GENETIC DISSECTION OF *DTH1.1*, A TRANSGRESSIVE FLOWERING TIME QTL, EXPRESSION AND DIVERSITY ANALYSIS OF FLOWERING TIME GENES IN RICE

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Luis Fernando Maas Molina

February 2010

GENETIC DISSECTION OF *DTH1.1*, A TRANSGRESSIVE FLOWERING TIME QTL, EXPRESSION AND DIVERSITY ANALYSIS OF FLOWERING TIME GENES IN RICE

Luis Ferrnando Maas Molina, Ph. D.

Cornell University 2010

Dth1.1 is a transgressive early flowering QTL that lies on the short-arm of chromosome 1. It was detected from an advanced backcross population of tropical japonica cv. Jefferson and *O. rufipogon*. Underlining *dth1.1* were large number of flowering time genes including *OsGI*, *FT-L8*, *OsSOC1*, *OsEMF1* and *PNZIP* that are components of the photoperiod pathway in rice. Genetic dissection of *dth1.1* was carryout through marker assisted selection developing a eight sub-introgression lines (SIL) that carry single or combination of candidate gene introgressions. SILs and controls were evaluated for flowering time under both short- and long-day growing conditions. Under short day lengths, lines with introgressions carrying combinations of linked flowering time genes (GI/SOC1, SOC1/FT-L8, GI/ SOC1/FT-L8 or EMF1/PNZIP) from the late parent, *O. rufipogon*, flowered earlier than the recurrent parent, Jefferson, while recombinant lines carrying smaller introgressions marked by the presence of *GI*, *SOC1, EMF1* or *PNZIP* alone no longer flowered early. Under long day length, lines carrying *SOC1/FT-L8, SOC1* or *PNZIP* flowered early, while those carrying *GI* or *EMF1* delayed flowering. A preliminary yield evaluation indicated that the transgressive early flowering observed in several of the SILs was also associated with a measurable and positive effect on yield. To understand underlying mechanism driving the transgressive early flowering observed in early

SILs, DNA sequence variation and expression analysis was conducted in four flowering time genes (*OsGI*, *HD1*, *HD3A* and *RFT1*). Expression levels of *HD3A* and to a lesser degree *RFT1* were predictive of flowering time, with higher *HD3A* mRNA levels associated with transgressive early flowering. These observations provide support for a flowering-time model whereby increases in the expression of *HD3A/RFT1* in chromosome 6 are caused by trans-acting factors located in the *dth1.1a* QTL region on chromosome 1. In our materials, *O. rufipogon* alleles across the *dth1.1a* QTL combined with Jefferson alleles in the *HD1-HD3A-RTF1* region on chromosome 6 are necessary to drive transgressive early flowering. The study laid the foundation for further study of photoperiod control transgressive flowering by looking at other *HD3A* induction pathways independently from *HD1* such as, *OsSOC1* and *EHD1*.

BIOGRAPHICAL SKETCH

Luis Fernando Maas Molina was born on January 31st, 1978 in Tegucigalpa, Honduras. He attended Hillcrest High School in Tegucigalpa, Honduras were he received a bilingual education and first became interested in biology after a series of class projects that involved experimentation with plants and dissections of animals. He then enrolled in Zamorano Pan American School of Agriculture in Zamorano, Honduras receiving an Agronomist degree in 1998 and a BSc in Agronomy with a double major in Horticulture and Food Science in 1999. While at Zamorano, he completed an undergraduate thesis under Dr. José Cojulun and Dr. José Miselem entitled "Analysis of thirty two factorial systems of ready-to-eat lettuce salads" were he was able to combine his interest in horticultural production systems and maintenance of food safety. In January 2000 he joined the Tropical Research Department of Standard Fruit Company a subsidiary of Dole Inc. as an applied research assistant in cantaloupe production farms in Choluteca, Honduras. This working experience was a turning point in his life. He was intellectually challenged by his supervisor, Dr. Francisco Gomez, the Director of the R&D Department and a mentor, who encouraged and supported him with the decision of pursuing a Master degree study abroad.

Luis was admitted in the University of Arkansas were he obtained a MSc degree in Horticulture and Molecular Biology in 2003. His thesis entitled "Development of Glufosinate-Resistant Spinach" involved the genetic transformation of spinach via agrobacterium mediated and site-specific introgression system Cre-LOX under the supervision of Dr. Teddy Morelock and Dr. Vihba Srivastava. It was here were he developed a keen interest in applied plant breeding and the use of molecular tools for the improvement of crops. He began his Ph.D. studies in the fall of 2003 at Cornell University in the Department of Plant Breeding and Genetics. He Joined the Dr. Susan

McCouch Lab in 2005 were he has undertaken a series of studies in flowering time in rice that has expanded his knowledge in conventional and molecular plant breeding.

Luis is married to Maria Guadalupe Salas Fernandez and has one daughter, Maria Luisa Maas, born on January 13th, 2001 in Coban, Guatemala and one Son, Felipe Eduardo Maas, born on April 28th, 2009

This dissertation is dedicated to my parents, Suyapa and Roberto, who sacrifice much throughout my life to supported my scientific career, to my wife Maria Guadalupe, with whom I share the passion for plant breeding and fight against hunger, to my daughter and son for showing me the true meaning of unconditional love and to the loving memory of my grandmother Carmen Molina and Marie Lavallard.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my major advisor, Dr. Susan McCouch for allowing me to join her program on very challenging terms and her continual support, encouragement and mentoring. I wish to thank Dr. William Coffman and Dr. Thomas Brutnell for taking the time to serve on my committee. I would also like to thank Dr. Martha Mutschler for the valuable training and experiences I received during my time in her lab. To the present and past members of the McCouch lab I am very thankful for your assistance and most importantly your friendship in particular to Fumio "Gen" Onishi, for the lessons pass to me during our interactions in greenhouse and for taking great care of the plants; to Sandy and Jenny for keeping the lab running smoothly and for always being willing to deal with all my requests; to Lois for her steady hand and guidance to get over all administrative hurdles and to everyone else for their friendship and helpfulness: Adam, Anjali, Megan, Jeremy, Michael, Mande, Noel, Chih-wei, Ize, Liang Xia, Manna and Dean. I am eternally grateful to Ms. Marie Lavallard for encouraging me to apply to Cornell University and the financial support provided for my studies in the form of a fellowship.

I would like to acknowledge the Department of Plant Breeding at Cornell University and to all professors, who believed in my capacity as a graduate student, gave me the opportunity to be part of an internationally recognized Department and enriched my Ph.D. with their scientific experience. My recognition also goes to all staff in particular Marggy Vangeli and Cynda Farnham. To my colleagues and fellow students, who helped me and supported me all the way during these five years of study.

Finally, to my parents Suyapa and Roberto and siblings Jessie, Juan, Roberto,

Carlos, Consuelo and Waleska for supporting my dreams and decisions throughout my career. They made me feel accompanied and loved throughout these very challenging times, even though it implied a physical separation during the last eight years. To my wife, Maria Guadalupe for her constant companionship, support, encouragement and belief in my abilities.

TABLE OF CONTENTS

Biographical sketch iii

Dedication v

Acknowledgements vi

Table of contents ix

List of figures x

List of tables xi

List of abbreviations xiii

Chapter 1. General Introduction 1

References 17

Chapter 2. Dissection of a QTL reveals adaptive, interacting gene complex associated with transgressive variation for flowering time and yield in rice.

Abstract 22

Introduction 23

Materials and Methods 26

Results 32

Discussion 49

Appendices 57

References 64

Chapter 3. Analysis of expression and nucleotide diversity of *OsGI*, *HD1*, *HD3A* and *RFT1* provides new insights into the genetic basis of

transgressive flowering time in rice.

Abstract 69
Introduction 70
Materials and Methods 76
Results 83
Discussion 97
Appendices 106
References 109

Chapter 4. Variation for flowering time among forty six diverse landraces of rice.

Abstract 117
Introduction 118
Materials and Methods 122
Results 123
Discussion 132
Appendices 136
References 141

Chapter 5. Conclusions and future perspectives of flowering time studies in rice.

Conclusions 149
Future Perspectives 151
Appendix 155

LIST OF FIGURES

Figure 1.1. Flowering time pathways networks in *Arabidopsis thaliana.* 4

Figure 1.2. Comparative genetics of photoperiod pathway in *Arabidopsis* and rice. 7

Figure 1.3. Regulation of flowering determinant gene *FLOWERING TIME LOCUS T* orthologs in *Arabidopsis*, Rice and *Pharbitis nil.* 13

Figure 2.1. Genetic scheme and pedigree showing development of Introgression Lines (SILs) and controls. 34

Figure 2.2. Graphical genotypes of the SILs and controls showing regions of *O. rufipogon* introgression 37

Figure 2.3. Bar graph showing days to flowering of SILs and controls in relation to recurrent parent, Jefferson. 41

Figure 2.4. Line graph showing average yield ($kg.ha^{-1}$) of SILs and controls in Beaumont, TX and Stuttgart, AR. 46

Figure 3.1. Essential components in the photoperiod pathway of rice. 71

Figure 3.2. (A) Graphical genotypes and allelic composition of lines evaluated. (B) Flowering time comparison of lines under short-day conditions. 78

Figure 3.3. Expression analysis of flowering time genes in parental lines Jefferson and *O. rufipogon* at four developmental stages. 85

Figure 3.4. Diurnal rhythms of (A) *OsGI*, (B) *HD1*, (C) *HD3A* and (D) *RFT1* in parental lines Jefferson and *O. rufipogon* under SD conditions. 88

Figure 3.5. Expression levels of *OsGI* (A), *HD1* (B), *HD3A* (C) and *RFT1* (D) in lines evaluated under SD conditions. 90

Figure 3.6. Model of photoperiod dependant flowering pathway in parental lines Jefferson, *O. rufipogon* and an early-flowering SIL. 99

Figure 4.1. Days to flowering comparison of 46 diverse rice lines under growth

chamber conditions. 127

Figure 4.2. (A) Days flowering distribution within each sub-population (B) Sub-population comparison under SD (dark blue bars) and LD (red bars) conditions. 129

Figure 4.3. Photoperiod sensitivity comparison of 46 diverse rice lines under growth chamber conditions. 133

LIST OF TABLES

Table 1.1. Photoperiod gene network comparison between *Arabidopsis* and rice. 10

Table 2.1. Flowering time gene allele specific markers. 31

Table 2.2. Dunnet's multiple means test (P<0.05) of days to flowering in relation to recurrent parent Jefferson. 43

Table 2.3. (A) ANOVA of Days to flowering (DTF) under growth chambers (B) ANOVA of Days to flowering (DTF) under field conditions (C) ANOVA for yield under field conditions. 44

Table 2.4 Yield performance ($kg.ha^{-1}$) comparison of SILs and controls using Dunnet's multiple means test (P<0.05). 48

Table 3.1. Sequence variations detected in coding and non-coding regions of flowering time genes *OsGI*, *HD1*, *HD3A* and *RFT1*. 94

Table 3.2. Functional classification of genes identified in SILs based on Gene Ontology (GO). 95

Table 4.1. Days to flowering time of 46 diverse rice lines under short day and long day photoperiod. 124

Table 4.2. Mean comparison (P<0.001) of days to flowering and Photoperiod sensitivity in *O.sativa* sub-populations under short and long day. 132

LIST OF APPENDICES

Appendix Table 2.1 List of INDEL and SSR markers, chromosomal locations, primer sequences and expected size. 62

Appendix Figure 2.1 Genomic distribution of homologous flowering time genes in *Arabidopsis* and *O. sativa*. 67

Appendix Table 2.2 Summary of agronomic trait means in SILs and parental lines. 68

Appendix Figure 3.1 Normalization gene comparison in parental lines Jefferson and *O. rufipogon*. 114

Appendix Table 3.1a List of flowering time allele specific, normalization and RT-PCR primers. 115

Appendix Table 3.1b List of primers for flowering time gene sequencing. 116

Appendix Table 4.1. Mean comparison of 46 diverse lines under short day photoperiod. 144

Appendix Table 4.2 Mean comparison of 46 diverse rice lines under long day photoperiod. 146

Appendix Table 4.3 ANOVA of flowering time in 46 diverse rice lines under growth chambers. 148

Appendix Table 5.1. Plant materials for gene expression analysis and selection of missing SILs. 157

Appendix Table 2. Primer sets for gene expression analysis. 158

Appendix Figure 5.1. (A) Breeding scheme and graphical genotype of Selected SILs. 160

LIST OF ABBREVIATIONS

ANOVA	Analysis of variance
ACT	Actin
ACT1	Actin 1
AR	Arkansas
CGS	Candidate gene specific marker
CRBD	Complete randomized block design
CSSL	Chromosome segment substitution line
CO	CONSTANS
DAG	Days after germination
DAG	Days after germination
DN	Day neutral
DTF	Days to flowering
DTH1.1	Days to heading 1.1
EMF1	EMBRYONIC FLOWER 1
EP	Expressed protein
FD	FLOWERING LOCUS D
FT	FLOWERING LOCUS T
FT-L8	FLOWERING LOCUS T LIKE 8
GI	GIGANTEA
GLM	General linear model
HD1	HEADING DATE 1
HD3A	HEADING DATE 3A
LD	Long day
MAS	Marker assisted selection

LSM	Least square model
NIL	Near isogenic line
PNZIP	*Pharbitis nil* LEU ZIPPER
PS	Photoperiod sensitivity
QTL	Quantitative trait loci
RCBD	Randomized complete block design
REML	Restricted mazimum likelihood
RFLP	Restriction fragment length polymorphism
RT-PCR	Real time polymerase chain reaction
SD	Short day
SIL	Single introgression line
SNP	Single nucleotide polymorphism
SOC1	SUPPRESOR OF OVEREXPRESSION OF CONSTANS 1
SSR	Simple sequence repeat
TX	Texas
UBQ	UBIQUITIN

CHAPTER 1

INTRODUCTION TO FLOWERING TIME STUDIES

The life cycle of plants can be divided into five distinctive stages: germination, vegetative growth, flowering, seed maturation and seed dispersal. The timing of flowering is one of the most important decisions during plant development. It directly impacts reproduction and is designed to coincide with optimum resource availability.

Plants have elaborate mechanisms for perceiving the direction, wavelength composition and duration of the light, temperature and other environmental cues that affect development (Garner and Allard, 1920). Most environmental factors vary from year to year, but day-length changes follow a predictable pattern that has allowed plants to evolve mechanisms that anticipate seasonal changes by integrating day-length information into their developmental programs. Garner and Allard (1920) demonstrated that many plants flower in response to changes in day-length. They explored day-length dependant flowering responses in tobacco and other species, allowing them to introduce the terms 'photoperiod' (a daily recurrent pattern of light and dark periods) and 'photoperiodism' (response or capacity to respond to photoperiod) (reviewed by Salisbury, 1985). They went on to classify plants according to photoperiodic responses as short-day (SD), long-day (LD) or day-neutral (DN) according to their response to day-length. SD plants are induced to flower when day-length is shorter than a specific critical day-length that varies between and within species. In contrast LD plants flower when the duration of the day-length exceeds this critical day-length (Salisbury, 1985; Putteril et al. 2004).

Mature leaves were determined to be the day-length sensing organs and a graft-transmissible substance was identified that was able to promote the switch from

the vegetative to the reproductive stage in species with different photoperiodic responses (Lang *et al.* 1977; Zeevart, 1976). The flowering substance was named 'florigen' and it was thought to move through the phloem from the leaves to the distal bud where it triggers the mechanism responsible for the transition from the vegetative to the reproductive stage (Chailakhyan, 1968; Pennazio, 2004).

Two models have been proposed to explain the integration of environmental and internal cues in the regulation of flowering by day-length. The *external coincidence model* proposes that the circadian clock generates a rhythm with an approximate 24-h period that controls flowering and is sensitive to light at a particular phase of the rhythm (Bunning, 1976; Imaizumi and Kay, 2006). The *internal coincidence model* proposes that the floral response occurs under conditions in which two differentially entrained rhythms are brought into the same phase under day-lengths that promote flowering (Imaizumi and Kay, 2006; Kobashi and Weigel, 2007). Among them, the *external coincidence model* is currently the most consistent with genetic evidence in plants (Yanovsky and Kay, 2003; Hayama and Coupland, 2004; Putterill et al. 2004). Erwin Bunning proposed the original hypothesis in 1936 based on studies of circadian and photoperiodic response of soybeans (Bunning, 1976; Saunders, 2005). The model was later expanded and refined in studies of insects with abnormal developmental programs (Saunders, 2005). In this model, light plays two crucial roles: (1) it resets the circadian clock, which is important for generating the daily oscillation of a key regulatory component with peak expression in the late afternoon; and (2) it regulates the activity of this component (Imaizuni and Kay, 2006). Flowering time will only be triggered when regulator levels above the threshold coincide with daylight, the external signal (Imaizuni and Kay, 2006).

Research in flowering time can be divided into two periods. The classical period is based on descriptive experiments (early 1920s) focusing on the contribution

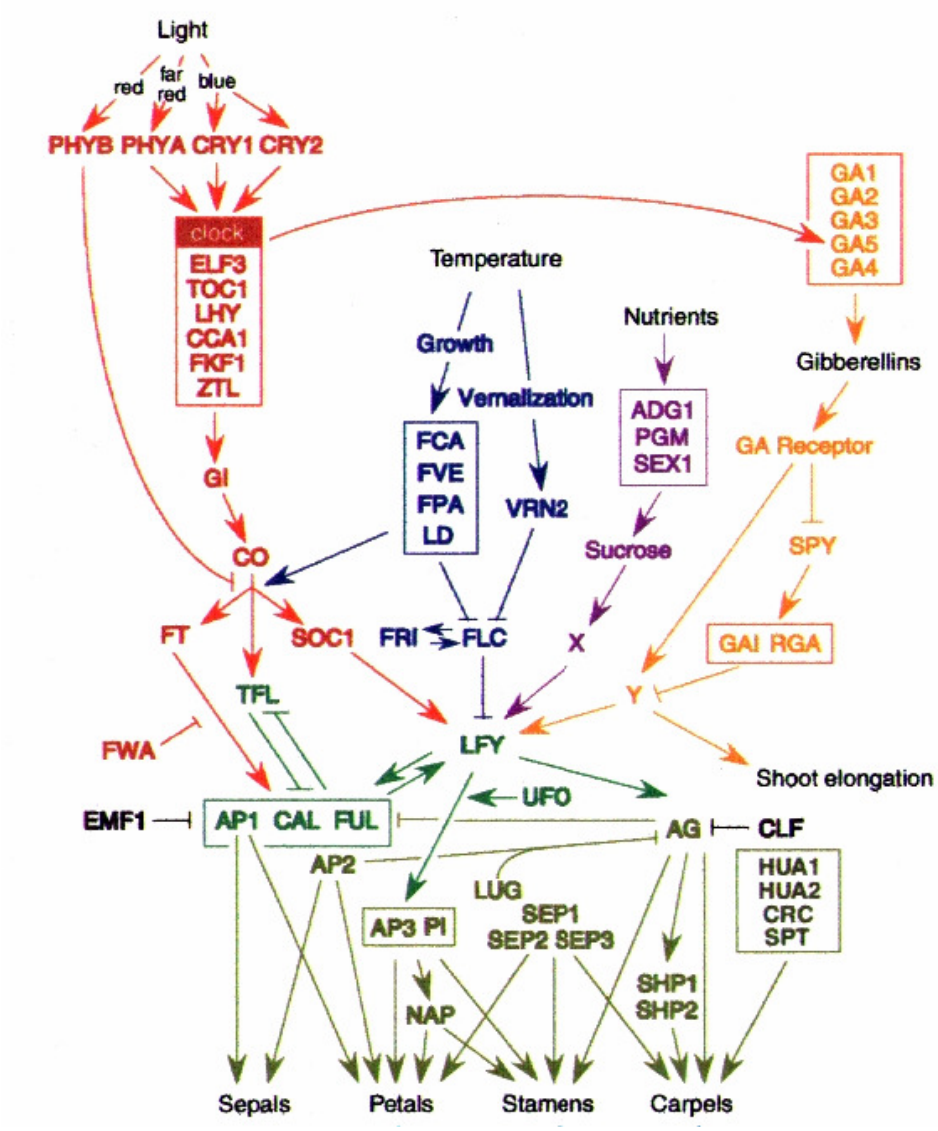
of day-length to the control of flowering time in plants that have different requirements for floral induction (Thomas and Vince-Prue, 1997). The second period is marked by genetic analysis of flowering mutants and rigorous hypothesis testing (Korneef et al. 1991). These experiments lead to a wealth of knowledge about the basic genetic regulation of flowering time in plants and paved the way for comparative studies in *Arabidopsis thaliana* and rice (*Oryza sativa*) and will likely serve as the foundation for broad sequence-based comparisons of many plant species.

Modern view of flowering time studies in model plants

The regulation of flowering time in plants has been most thoroughly studied in the model long day plant *Arabidopsis* where at least four distinct genetic pathways contribute to the transition from the vegetative to the reproductive stage (Fig. 1.1). They are: (1) the photoperiod promotion pathway, (2) the constitutive or autonomous pathway, (3) the vernalization pathway, and (4) the gibberellic acid promotion pathway (Blazquez, 2000; Mouradov et al. 2002; Simpson and Dean 2002; Yanovsky and Kay 2003; Putterill et al. 2004). Current data show that floral development is a repressible process when flowering time genes in these pathways are expressed. This suggests that late-flowering is the ancestral character in both rice and *Arabidopsis* (Komeda, 2004).

The photoperiodic pathway involves photoreceptors (including phytocromes and cryptochromes), the circadian clock and several output pathways (Mouradov et al. 2002; Cremer and Coupland, 2003; Tsuji et al. 2008). Comparative genetic studies of the photoperiodic control of flowering time in the model short day plants rice and Japanese morning glory, the long day plant *Arabidopsis*, and in crops such as wheat, maize, soybean and poplar, have determined that genes in the photoperiod pathway are highly conserved across the plant kingdom (Hayama and Coupland, 2004; Izawa et al. 2003; Izawa, 2007).

Figure 1.1 Flowering time pathways networks in *Arabidopsis thaliana*: Photoperiod (red), Vernalization (blue), Autonomous (purple) and Giberellic acid (orange). Meristem identity genes are in green (from Blazquez, 2000).

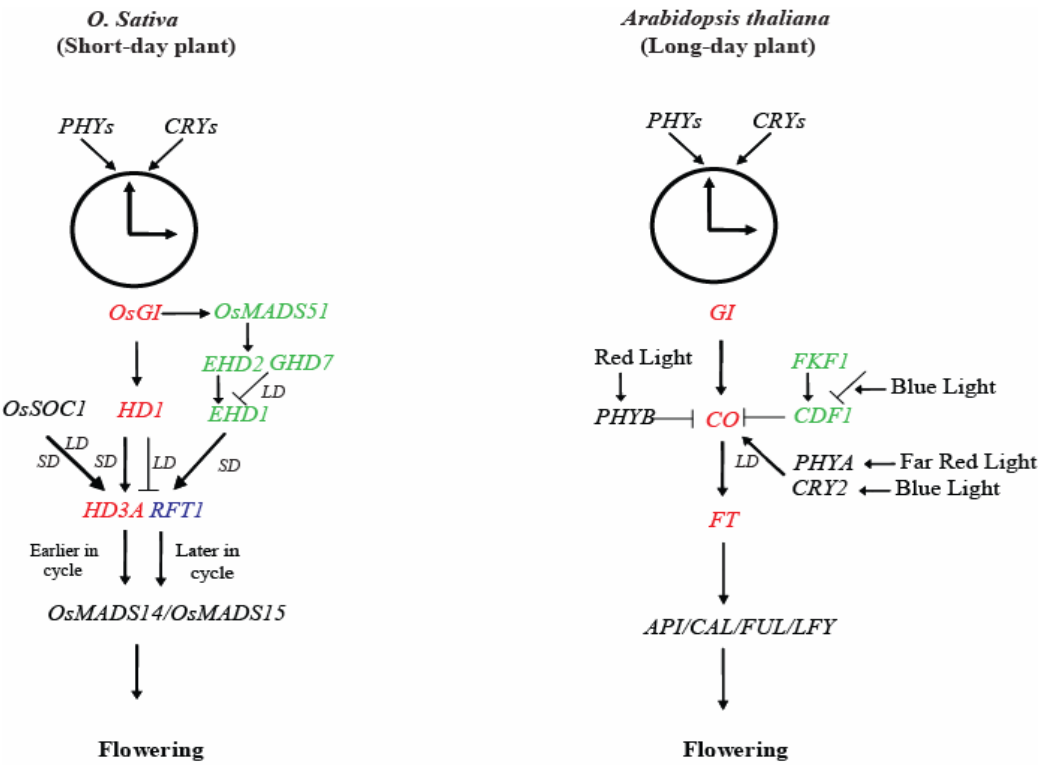

Light
red
far red
blue
PHYB PHYA CRY1 CRY2
clock
ELF3
TOC1
LHY
CCA1
FKF1
ZTL
GI
CO
FT
SOC1
TFL
FWA
Temperature
Growth
Vernalization
FCA
FVE
FPA
LD
VRN2
FRI
FLC
Nutrients
ADG1
PGM
SEX1
Sucrose
X
GA1
GA2
GA3
GA5
GA4
Gibberellins
GA Receptor
SPY
GAI RGA
Y
LFY
UFO
Shoot elongation
EMF1
AP1 CAL FUL
AG
CLF
AP2
LUG
SEP1
SEP2 SEP3
AP3 PI
HUA1
HUA2
CRC
SPT
SHP1
SHP2
NAP
Sepals
Petals
Stamens
Carpels

These studies suggest that most of the phenotypic variation between rice and *Arabidopsis* is mediated by differential regulation of genes such as *GIGANTEA (GI)*, *CONSTANS (CO)* and *FLOWERING LOCUS T (FL-T)* (Laurie, 1997; Cremer and Coupland, 2003; Hayama and Coupland 2004; Putterill et al. 2004) (Fig. 1.2). The interaction among the components of the photoperiodic output pathway suggests that *GI* is required to increase CO transcription but that post-transcriptional levels of CO are determined by the antagonistic action of *PHYTOCHROME A* (*PHYA*), *CRYPTCHROME 2* (*CRY2*) and *PHYTOCHROME B* (*PHYB*). *CO* is also critical in mediating the interaction between circadian rhythms and light signaling which occurs via *CO* transcription and CO protein stability.

The CO protein then activates the transcription of *FLOWERING TIME LOCUS T* (*FT*) and *SUPPRESOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*). *FT* and *SOC1* act as the integrators of the photoperiod, giberrelic and vernalization pathways and the induction of *FT* above a threshold level promotes the expression of meristem identity genes that, in turn, cause the apical meristem to switch from vegetative to reproductive growth (Kobayashi and Weigel, 2007). The FT protein moves through the phloem and interacts with the transcription factor, *FLOWERING TIME LOCUS D* (*FD*), which is expressed in the flanks of the shoot apical meristem where the flower primordia is initiated (Abe, M et al. 2005; Tamaki et al. 2007; Turck et al. 2008). In the presence of *FT*, *FD* is able to induce a number of important target genes including *LEAFY*, *APETALA1* and *FRUITFULL*.

Previous studies have identified putatively orthologous flowering time genes in rice for *GI*, *CO* and *FT* (Izawa et al. 2003) (Fig. 1.2 and Table 1.1). *HEADING DATE 1* (*HD1*), *HEADING DATE* 3a (*HD3A*) and *HEADING DATE 6* (*HD6*) have been cloned and found to encode proteins similar to CO, FT and the α-subunit of casein kinase 2, respectively (Yano et al. 2000; Takahashi et al. 2001; Kojima et al. 2002).

Figure 1.2 Comparison of the photoperiod pathway in *Arabidopsis* and rice. A core set of genes: *GI:OsGI*, *CO:HD1* and *FT:HD3A/RFT1* are conserved but the function and expression control of the genes differentiate long day *Arabidopsis* and facultative short day rice. Two other sources of variation are induction flowering pathway unique to each species such as *OsMADS51:EHD2:EHD1* and the duplication of *FT:HD3A* genes in rice including *RFT1* that have been shown to be essential for flowering in rice.


O. Sativa
(Short-day plant)
PHYs
CRYs
OsGI
OsMADS51
EHD2 GHD7
LD
OsSOC1
HD1
EHD1
LD
SD
SD
LD
SD
HD3A RFT1
Earlier in cycle
Later in cycle
OsMADS14/OsMADS15
Flowering
Arabidopsis thaliana
(Long-day plant)
PHYs
CRYs
GI
Red Light
FKF1
Blue Light
PHYB
CO
CDF1
LD
PHYA
Far Red Light
CRY2
Blue Light
FT
API/CAL/FUL/LFY
Flowering

OsGI was isolated by differential display and mRNA expression was found to be circadian-controlled with a temporal expression pattern that was very similar to that of *AtGI* under both SD and LD conditions. Over expression and RNAi silencing experiments have shown that *OsGI* inhibits flowering in rice under short and long-days, suggesting a reversal in the regulatory function of the gene between *Arabidopsis* and rice (Hayama et al. 2003).

The rice *HD1* gene is required for the suppression of flowering under LD and for the promotion of flowering under SD conditions. Furthermore, the *HD3A* gene was shown to be an activator of flowering in rice. Under LDs, the *HD1* protein, which is expressed at the end of the day, is activated by phytochrome and inhibits flowering by shutting down *HD3A* expression (Hayama et al. 2003). In contrast, under SDs, *HD1* is expressed only during the night when phytochrome is hypothesized to be inactivate, and this allows *HD1* to induce *HD3A* expression and promote flowering.

Another flowering-time gene of rice, *EARLY HEADING DATE 1* (*EHD1*), has no obvious orthologue in *A. thaliana* (Table 1.1 and Figure 1.2; Doi et al. 2004). *EHD1,* encodes a B-type response regulator that promotes floral transition preferentially under SD conditions, even in the absence of functional alleles of *HD1*. Therefore, *HD1* and E*HD1* function redundantly under SD but antagonistically under LD conditions (Fig. 1.2). This antagonistic action makes it possible for rice to flower even under long day conditions. Rice plants carrying a functional *HD1* and a non-functional *EHD1* did not flower under long day conditions, even after 180 d (Doi et al. 2004). Extensive expression analysis revealed that *EHD1* is preferentially expressed under short day conditions and acts upstream of *FT* orthologues such as *HD3A* (Doi et al. 2004).

These results clearly indicate that two independent floral pathways, the evolutionally conserved *HD1* pathway and the unique *EHD1* pathway, integrate

environmental photoperiod signals into the expression of *FT* orthologues (such as *HD3A*) and make rice a short day plant (Izawa et al.2003; Izawa, 2007).

Table 1.1 Photoperiod gene network comparison between *Arabidopsis* and rice. Empty cells indicate no corresponding homolog between both species.

Category	Gene Name	*Arabidopsis*	Rice
Photoreceptor	PHYTOCHROME A	PHYA	PHYA
	PHYTOCHROME B	PHYB	PHYB
	PHYTOCHROME C	PHYC	PHYC
	PHYTOCHROME D	PHYD	
	CRYOTOCHROME 1	CRY1	CRY1
	CRYPTOCHROME 2	CRY2	CRY2
Circadian Clock	CIRCADIAN CLOCK ASSOCIATED	CCA1	
	LATE ELONGATED HYPOCOTYL	LHY	
	GIGANTEA	GI	
	TIMING OF CAB EXPRESSION 1	TOC1/APRR1	OsPRR
	ARABIDOPSIS PSEUDO-RESPONSE REGULATOR	APRR3	
		APRR5	
		APRR7	
		APRR9	
	EARLY FLOWERING 3	ELF3	
	ZEITULE	ZTL	
	LOV KELCH PROTEIN 2	LKP2	
Circadian Clock Mediator	FLAVIN BINDING KELCH REPEAT 1	FKF1	OsFKF1
	GIGANTEA	GI	OsGI
	CONSTANS	CO	HD1/SE1
		HYI	SE5
			HD6
			HD9
Floral Pathway Integrator	FLOWERING LOCUS T	FT	HD3A/FT-L2
			FTL/FT-L1
			RFT1/FT-L3
			OsFT/FT-L4
			FT-L5
			FT-L6
			FT-L7
			FT-L8
			FT-L9

Table 1.1 (continue)

		FT-L10
		FT-L11
		FT-L12
SUPPRESOR OF OVEREXPRESSION OF CONTANS 1	SOC1	OsSOC1/OsMADS50
MADS-BOX 51		OsMADS51
LEAFY	LFY	RFL
FLOWERING LOCUS D	FD	
EARLY HEADING DATE 1		EHD1
EARLY HEADING DATE 2		EHD2
GRAIN HEIGHT DAYS TO HEADING 7		GHD7

These studies confirm the presence of a conserved photoperiod pathway between *Arabidopsis,* rice, and *Pharbitis nil* while at the same time providing clues to the reversal of gene function that leads to the difference between short day and long day plants, by the presence of unique induction pathways in each species and the duplication of the FT gene family in rice (Fig. 1.3 and Table 1.1).

Flowering time as a breeding tool

Manipulation of natural variation in flowering time by plant breeders has expanded the area of cultivation of agronomic species from their original center of origin to many diverse environmental conditions (Izawa, 2007). At the same time it has allowed for significant increases in planting densities by manipulating height and leaves angles, increasing the productivity per unit area and reducing the amount of land necessary to feed an ever growing demand for food.

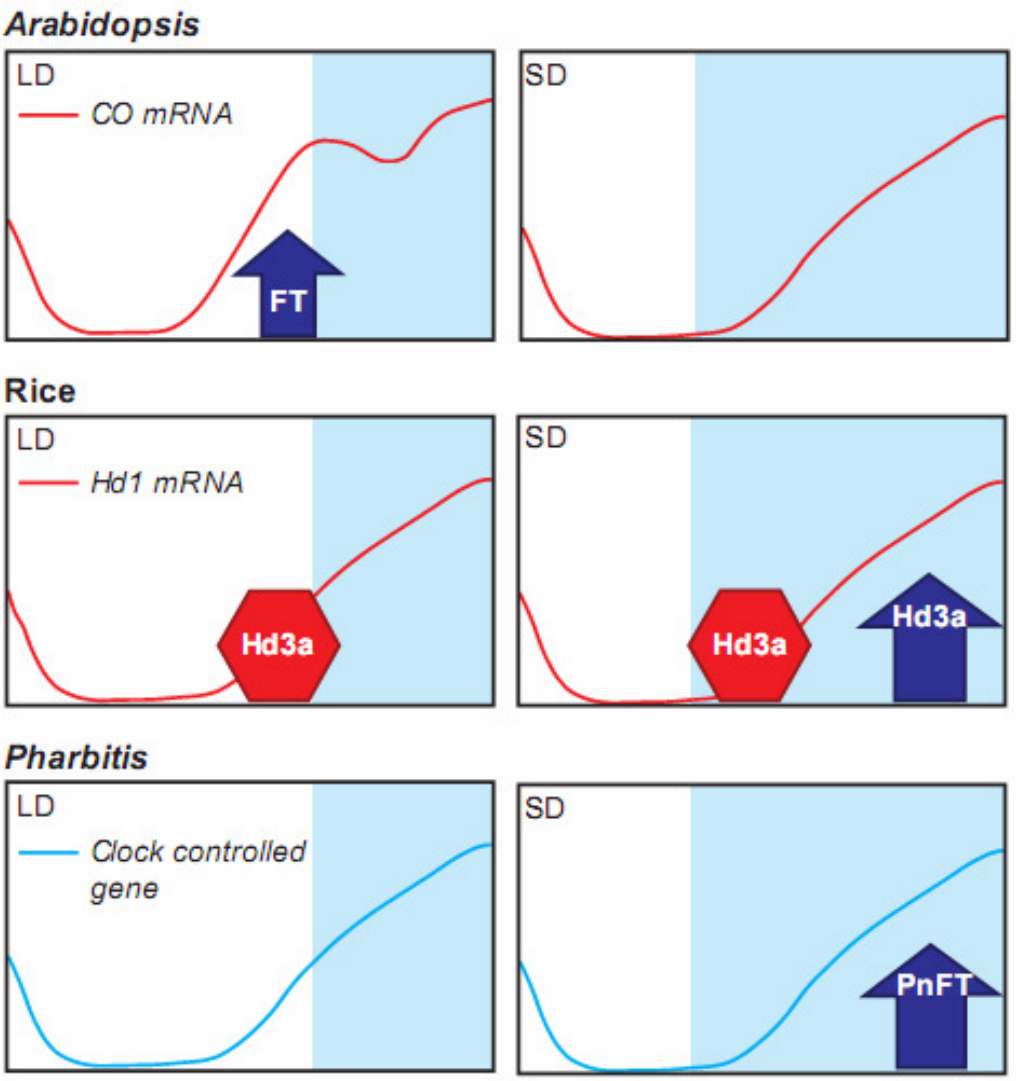
Unlike laboratory-generated variation, adaptive natural variation results from the long process of natural selection whereby mutations are screened according to their phenotypic global effect (Roux et al, 2006). Studies of natural variation reveal a complex interactive process and offer many new insights about how flowering time is

controlled in different species, often via unexpected changes in genes of known function such as *HD1* in rice and *FT-L2* in poplar (Cremer and Coupland, 2003; Izawa, 2007).

Cultivated Asian rice, *O. sativa*, is derived from the wild rice, *Oryza rufipogon* (Khush, 1997; Vitte et al. 2004). *O. rufipogon* is broadly distributed around Indochina, from southern China to eastern India, mainly at tropical latitudes (Londo et al. 2006). Its northern limit is currently around 28^{o} N. Palaeo-botanical studies revealed that the northern limit of the ancestral wild rice was near the Yangtze River basin in China at around 31^{o} N several thousand years ago (Cao et al. 2006; Lee et al. 2007). By contrast, cultivated rice is distributed widely; its northern limit is currently around 45^{o} N (Izawa, 2007). This expansion has been made possible by domestication and breeding over the past several thousand years. It is apparent that strong artificial selection has adapted rice to these broader areas. This northward expansion of rice into more temperate regions is largely dependent on changes in the flowering time of cultivars, in addition to the acquisition of cold-tolerance traits (Izawa, 2007).
At higher latitudes, early flowering and reduced photoperiod-sensitivity are essential to producing a harvest before the approaching cold weather makes plants sterile (Izawa, 2007). Thus, there is expected to be an association between the northward expansion of rice and natural variation in flowering-time genes.

In fact, it is well known that many rice cultivars exhibit a latitudinal cline in flowering time. This cline has a simple explanation. At higher latitudes, the periods available for flower formation, meiosis in pollen development, and embryogenesis become limited, so critical control of flowering time is essential for rice cultivation (Izawa, 2007). The domestication and breeding of rice for northern regions might have included steps that changed flowering time.

Figure 1.3 Regulation of the flowering determinant gene *FLOWERING TIME LOCUS T* orthologs in *Arabidopsis*, *O. sativa* and *Pharbitis nil* under long-day (LD) and short-day (SD) conditions (from Turck et al. 2008).


Arabidopsis
LD
CO mRNA
FT
SD
Rice
LD
Hd1 mRNA
Hd3a
SD
Hd3a
Hd3a
Pharbitis
LD
Clock controlled gene
SD
PnFT

Thus, the adaptation of rice cultivars to local regions at distinct latitudes can be partly explained by three types of photoperiodic control of flowering as seen in *temperate japonica* rice cultivars: short day floral promotion with redundant actions of two flower-promoting pathways by *HD1* and *EHD1*; long day floral repression with antagonistic actions by *HD1* and *EHD1*; and early flowering due to de-repression of *HD3A* expression by unknown mechanisms (Izawa, 2007). Selection on variations in flowering time regulation, at least partly explain how rice became adapted to such a broad range of environments during the process of domestication and breeding.

This dissertation characterizes the genetic basis of transgressive variation for flowering time associated with QTL *dth1.1* (*days to heading 1.1*) located on the short arm of chromosome 1 using near isogenic lines (NILs) derived from an advanced backcross population derived from *Oryza sativa* cv Jefferson (early flowering) as the recurrent parent, and the late-flowering wild rice relative, *Oryza rufipogon,* as the donor parent.

In Chapter Two I dissect the *dth1.1* QTL into five sub-introgression lines (SILs) and ask how recombination within a single defined complex QTL is capable of generating transgressive early flowering in rice. The focus was on five candidate flowering time genes that are loosely linked within the *dth1.1* interval: *GIGANTEA* (*OsGI*), *SUPRESSOR OF EVEREXPRESSION OF CONSTANS 1* (*OsSOC1*), *FLOWERING LOCUS T-LIKE 8* (*FT-L8*), *EMBRYONIC FLOWERING 1* (*OsEMF1*) and *Pharbitis nil LEU ZIPPER* (*PNZIP*). SILs containing one or more candidate genes were evaluated in controlled environments and under field conditions to determine the effect of each *O. rufipogon* introgression on flowering time and yield. By partitioning the observed variation, we were able to identify specific sub-introgressions that are responsible for the transgressive early flowering time and also contribute to an increase in harvestable yield. These SILs can be used as breeding lines to introduce

trait-enhancing alleles from *O. rufipogon* that will simultaneously help to broaden the genetic base of U.S. rice cultivars.

In Chapter Three, RT-PCR was use to characterize the expression of four flowering time genes in the different SILs to determine the expression levels and daily rhythms of *O. rufipogon* alleles at *GIGANTEA* (*OsGI*), *HEADING DATE 1* (*HD1*), *HEADING DATE 3a* (*HD3A*), and *RICE FLOWERING LOCUS T 1* (*RFT1*). Sequence from all four genes was obtained from both parental lines to infer how sequence variation (i.e., SNPs and INDELS, might affect the function of each gene. Information from these analyses provided the foundation for interpreting the genetic basis of transgressive variation for flowering time in the *dth1.1* SILs and furthered our understanding of the photoperiodic control of flowering time in rice. Because we introgressed genes from the wild relative *O. rufipogon,* our work provides a clear illustration of how introgressive hybridization between cultivated species and their wild relatives can give rise to dramatic shifts in phenotype and further, how recombination within a single QTL region may enhance the range of variation available to plant breeders.

Chapter Four lays the foundation for future association mapping studies of flowering time in rice using a candidate gene approach. Forty-five diverse rice lines from the five major *O. sativa* sub-populations were evaluated for flowering time and photosensitivity under controlled conditions and complete gene sequence from two flowering time genes, *GIGANTEA* (*OsGI*) and *HEADING DATE 1* (*HD1*), was obtained on the 45 lines. In this study, we start to bridge the gap between our understanding of flowering time at the molecular genetic level and our ability to use that information to make tangible progress in plant breeding through the exploitation of natural variation found in wild and cultivated species.

REFERENCES

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, and Araki T (2005) FD, a bZIP Protein Mediating Signals from the Floral Pathway Integrator FT at the Shoot Apex. Science 309: 1052

Blazquez M (2000) Flower development pathways. J Cell Sci 113: 3547-3548

Bunning E (1976) Circadian rhythms, light, and photoperiodism: A re-evaluation. Journal of Plant Research 92: 89-103

Cao ZH, Ding JL, Hu ZY, Knicker H, Kogel-Knabner I, Yang LZ, Yin R, Lin XG, Dong YH (2006) Ancient paddy soils from the Neolithic age in China's Yangtze River delta. Natur-wissenschaften 93: 232–236

Chailakhyan MK (1968) Internal factors of plant flowering. Annu. Rev. Plant Physiol. 19: 1-36

Cremer F and Coupland G (2003) Distinct photoperiodic responses are conferred by the same genetic pathway in *Arabidopsis* and rice. Trends in Plant Science 8: 405-407

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M and Yoshimura A (2004) *EHD1*, a B-type response regulator in rice, confers short day promotion of flowering and controls *FT-like* gene expression independently of *HD1*. Genes & Dev 18:926-936

Garner WW, Allard HA (1920) Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. J. Agric. Res. 18: 553-606

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K (2003) Adaptation of photoperiodic control pathways produces short day flowering in rice. Nature 422: 719-722

Hayama R and Coupland G (2004) The molecular basis of diversity in the photoperiodic flowering responses of *Arabidopsis* and rice. Plant Physiol. 135: 677-684

Imaizumi T and Kay S (2006) Photoperiodic control of flowering: not only by coincidence. Trends in plant science 11: 550-558

Izawa T, Takahashi Y and Yano M (2003) Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and *Arabidopsis*. Current Opinion in Plant Biology 6: 113-120

Izawa T (2007) Adaptation of flowering time by natural and artificial selection in *Arabidopsis* and rice Journal Exp Bot 58: 3091-3097

Khush GS (1997) Origin, dispersal, cultivation and variation of rice. Plant Molecular Biology 35: 25-34

Kobayashi Y, and Weigel D (2007) Move on up, it's time for change mobile signals controlling photoperiod-dependent flowering. Genes & Dev. 21: 2371-2384

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M (2002) HD3A, a rice ortholog of the *Arabidopsis* FT gene, promotes transition to flowering downstream of HD1 under short day conditions. Plant Cell Physiol 43: 1096-1105

Komeda Y (2004) Genetic regulation of time to flower in *Arabidopsis thaliana*. *Annual Review of Plant Biology* 55**:**521–535

Koornneef M, Hanhart CJ, van der Veen JH (1991) A genetic and physiological analysis of late-flowering mutants in *Arabidopsis thaliana.*. Mol. Gen. Genet. 229:57-66

Lang A *et al.* (1977) Promotion and inhibition of flower formation in a day neutral plant in grafts with a short-day plant and long-day plant. PNAS 74: 2412-2416

Laurie D (1997) Comparative genetics of flowering time. Plant Mol Bio 35: 167-177

Lee GA, Crawford GW, Liu L, Chen X (2007) Plants and people from the Early Neolithic to Shang periods in North China. PNAS 104: 1087–1092

Londo JP, Chian YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. PNAS 100: 10102-10107

Mouradov A, Cremer F, Coupland G (2002) Control of flowering time: interacting pathways as a basis for diversity. Plant Cell 14: S111-130

Pennazio S (2004) Florigen: an intriguing concept of plant biology. Riv. Biol. 97: 33-51

Putterill J, Laurie R and Macknight R (2004) It's time to flower: the genetic control of flowering time. Bioessays 26: 363-373

Roux F, Touzet P, Cuguen J and Le Corre V (2006) How to be early flowering: an evolutionary perspective. Trends in Plant Science 11: 375-381

Salisbury FB (1985) Photoperiodism. Horticultural Review 4: 66-105

Saunders DS (2005) Erwin Bunning and Tony Lees, two giants of chronobiology, and the problem of time measurements in insect photoperiodism. J, Insect Physiol. 51: 599-608

Simpson GG, Dean C (2002) *Arabidopsis*, the Rosetta stone of flowering time? Science 296: 285-289

Takahashi Y. *et al*. (2001) Hd6, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the a-subunit of protein kinase CK2 Proc. Natl. Acad. Sci. 98: 7922–7927

Tamaki S, Matsuo S, Wong HL, Yokoi S, and Shimamoto K (2007) HD3A Protein Is a

Mobile Flowering Signal in Rice. Science 316: 1033-1036

Thomas B , Vince-Prue D (1997) Photoperiodism in Plants, 2nd ed. (San Diego, CA: Academic Press)

Tsuji H, Tamaki S, Komiya R and Shimamoto K (2008) Florigen and the photoperiodic control of flowering in rice. Rice 1: 25-35

Turck F, Fornara F and Coupland G (2008) Regulation and identity of florigen: *Flowering Locus T* moves to center stage. Annu. Rev. Plant Biol. 59:573-594

Vitte C, Ishi T, Lamy F, Brar D, Panaud O (2004) Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L). Molecular Genetics and Genomics 272: 504-511

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) HD1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene CONSTANS. Plant Cell 12: 2473-2484

Yanovsky MJ and Kay SA (2003) Living by the calendar: how plants know when to flower. Nat. Rev. Mol. Cell. Biol. 4: 265-275

Zeevart JA (1976) Physiology of flower formation. Annu. Rev. Plant Physiol. 27: 321-348

CHAPTER 2

DISSECTION OF A QTL REVEALS ADAPTATIVE, INTERACTING GENE COMPLEX ASSOCIATED WITH TRANSGRESSIVE VARIATION FOR FLOWERING TIME AND YIELD IN RICE[1]

Abstract

A *days to heading* QTL (*dth1.1*) located on the short-arm of rice chromosome 1 was sub-divided into eight sub-introgression lines (SILs) to analyze the genetic basis of transgressive variation for flowering time. Each SIL contained one or more introgression(s) from *O. rufipogon* in the genetic background of the elite *O. sativa* cultivar, Jefferson. Each introgression was defined at high resolution using molecular markers and those in the *dth1.1* region were associated with the presence of one or more flowering time genes (*GI, SOC1, FT-L8, EMF1,* and *PNZIP*). SILs and controls were evaluated for flowering time under both short- and long-day growing conditions. Under short day lengths, lines with introgressions carrying combinations of linked flowering time genes (GI/SOC1, SOC1/FT-L8, GI/ SOC1/FT-L8 or EMF1/PNZIP) from the late parent, *O. rufipogon*, flowered earlier than the recurrent parent, Jefferson, while recombinant lines carrying smaller introgressions marked by the presence of *GI*, *SOC1, EMF1* or *PNZIP* alone no longer flowered early. Under long day length, lines carrying *SOC1/FT-L8, SOC1* or *PNZIP* flowered early, while those carrying *GI* or *EMF1* delayed flowering. Across all experiments and in the field, only SIL_SOC1/FT-L8 was consistently early. A preliminary yield evaluation indicated that the transgressive early flowering observed in several of the SILs was also

[1]Luis F. Maas, Anna McClung and Susan R. McCouch. 2009. Dissection of a QTL reveals adaptive, interacting gene complex associated with transgressive variation for flowering time in rice. TAG (*in press*)

associated with a measurable and positive effect on yield. These SILs represent a new source of variation that can be used in breeding programs to manipulate flowering time in rice cultivars without the reduction in yield that is often associated with early maturing phenotypes.

Introduction

The initiation of flowering is one of the most important transitions in the life cycle of a plant. It is regulated by a complex genetic network that integrates intrinsic developmental signals with environmental cues such as day length and temperature. The regulation of flowering time by both endogenous and exogenous signals ensures that flowering coincides with appropriate environmental conditions and leads to the successful reproduction of the species (Izawa et al. 2003; Turck et al. 2008)

Garner and Allard (1920) classified plants into three categories depending on their developmental response to specific day/night lengths: long day (LD), short day (SD) and day neutral (DN). SD plants are induced to flower when the day length is shorter than a particular duration, called the critical day length. In contrast LD plants flower when day length exceeds this critical value (inductive conditions). DN plants tend to flower similarly under both LD and SD. Obligate SD or LD plants require exposure to the critical day length before they flower, while facultative SD or LD plants are more likely to flower under the appropriate light/dark conditions, but will eventually flower regardless of day or night length.

In a previous study, eleven flowering time QTLs were mapped in an interspecific cross between the early flowering *Oryza sativa* cultivar, Jefferson, and a wild accession of *O. rufipogon* (Thomson et al. 2003). A QTL on the short arm of chromosome 1, *days to heading 1.1* (*dth1.1*; LOD=9.0; R^2 = 0.08- 0.15), promoted

early flowering in the recurrent parent (cv. Jefferson) despite the fact that *O. rufipogon* flowers much later than Jefferson (Thomson et al. 2006). The QTL was originally detected in BC_2F_2 populations in two field environments as well as under greenhouse conditions. The shape of the interval plot for *dth1.1* was very broad: for the field environment the QTL plot was significant (LOD > 3.0) across approximately 54 cM, while the plot for the greenhouse environment showed a significant QTL across 38 cM of the short arm of chromosome 1 (Thomson et al. 2006). A similar region had been previously associated with flowering time in other interspecific populations of rice (Cai and Morishima 2002; Doi et al. 1998; Kohn et al. 1997; Xiao et al. 1998).

Using substitution mapping, Thomson et al. (2006) dissected *dth1.1* into two, non-overlapping separate QTLs, *dth1.1a* and *dth1.1b,* and demonstrated that *O. rufipogon* alleles across both regions independently conferred earliness under both short-day (SD=10 hours) and long-day (LD=14 hours) lengths in the Jefferson genetic background. However, the lines used for substitution mapping retained several *O. rufipogon* introgressions on other chromosomes, leaving open the possibility that genes outside the *dth1.1* region may also have contributed to transgressive early flowering.

Six flowering time genes were identified within the *dth1.1* QTL interval, based on sequence similarity to known flowering time genes from *Arabidopsis* (*GIGANTEA (OsGI), FLOWERING TIME LOCUS T (OsFTL and OsFT-L8), SUPPRESSOR OF CONSTANS 1 (OsSOC1), EMBRYONIC FLOWER 1 (OsEMF1)* and from *Pharbitis nil* (*Leucine Zipper* (*OsPNZIP)*) (Thomson et al. 2006). The flowering time orthologs under *dth1.1* were all involved in the photoperiodic pathway. This pathway controls flowering in response to day length and involves photoreceptors (including phytocromes and cryptochromes), the circadian clock and several output pathways (Blazquez 2000; Izawa et al. 2002; Park et al. 1999). While these genes are all loosely

linked within the 38 cM QTL region on the short arm of rice chromosome 1, their homologous counterparts in *Arabidopsis* show no linkage and are distributed on 4 different chromosomes (Supplementary Fig. 2.1). Rice and *Arabidopsis* are estimated to have diverged from a common ancestor ~200 MYA, and the average size of conserved syntenic regions is <3 cM (Paterson et al. 2009; Salse et al. 2002). This suggests that the evolutionary process would have disrupted any extensive gene complexes that existed in a common ancestor,and raises interesting questions about the origin of the array of flowering time genes found on rice chromosome 1.

Comparative studies in *Arabidopsis*, rice and *Pharbitis nil* have demonstrated that there is a highly conserved network of genes involved in the photoperiod pathway in both LD and SD plants (Izawa et al. 2003; Kojima et al. 2002). The basis for the differential flowering responses to photoperiod derive from duplications and changes in the expression of gene family members (Chardon and Damerval 2005; Komiya et al. 2008), development of alternative flowering induction pathways that are unique to each species (Doi et al. 2004; Kim et al. 2007; Komiya et al. 2008; Matsubara et al. 2008; Tsuji et al. 2008) and a reversal in the function of at least three genes in rice and *Arabidopsis*, *GIGANTEA (GI*), *CONSTANS (CO* or *HD1*), and *FLOWERING TIME LOCUS T (FT* or *HD3A*) (Hayama and Coupland 2004; Hayama et al. 2003; Kojima et al. 2002; Putterill et al. 2004). In rice, *GI* resides within the cluster of flowering time candidates in the *dth1.1a* QTL region on rice chromosome 1, while *HD1* and *HD3A* are located within 6.4 Mb of each other on the short arm of chromosome 6.

Although tremendous progress has been achieved in understanding which genes are involved in the regulation of flowering time and how different genes and gene families interact at the molecular level, our understanding of how flowering time is regulated under field conditions and how allelic variation and copy number variation in natural populations affects the flowering response of our major crop plants

remains largely unknown. The introgression lines in our study provide an opportunity to investigate how allelic variation within the *dth1.1* QTL region contributes to variation in flowering time in the presence of either *O. rufipogon* or Jefferson DNA across the region on chromosome 6, known to contain a second cluster of flowering time genes.

The objectives of the present study were to (1) construct a set of sub-introgression lines (SILs) in the genetic background of the USA *tropical japonica* cultivar, Jefferson, each containing a well-defined introgression from *O. rufipogon* carrying one or more candidate genes for flowering time, (2) evaluate the SILs for flowering time under controlled conditions in both long and short days, and (3) evaluate the SILs for yield and flowering time under field conditions. We were interested in dissecting the *dth1.1* QTL to better understand the genetic basis of transgressive variation for flowering time and to determine whether creating novel haplotypes via recombination within the QTL region was capable of producing novel transgressive phenotypes that could be of use in plant improvement.

Materials and Methods

Plant Material: All of the introgression lines in this study are derived from BC_4F_4 families (Jefferson/*O. rufipogon* (IRGC #105491)////Jefferson) as described by Thomson et al. (2006). The term "SIL" refers to BC_4F_8 sub-introgression lines that each contained a defined introgression from *O. rufipogon* that sub-divides the original 38 cM *dth1.1* QTL target region. "Pre-SILs" are selected BC_4F_4 lines that served as initial materials from which the SILs were derived (Fig. 2.1). SILs were named according to the candidate gene(s) contained in each *O. rufipogon* introgression, and for ease of reading, will hereafter be referred to simply by the gene name, written

without italics.

Four pre-SILs (P10-92, P1-76, P6-78 and P9-70 in Fig. 2.1) were selected as the starting material for this study. Each contained a heterozygous introgression from *O. rufipogon* across the *dth1.1* region and demonstrated transgressive variation for early flowering under short days in the study by Thomson et al. (2006). When selfed, these lines gave rise to SILs with homozygous introgressions in the target region(s), heterozygous offspring, and "revertant" SIL controls containing no *O. rufipogon* DNA in the target regions. These revertants, P10-92R, P1-76R, P6-78R and P9-70R, were included as controls and enabled us to evaluate the phenotypic effect of spurious *O. rufipogon* introgressions remaining in the genetic background of the SILs. Lines P9-84, P4-12 and P15-62 were also included as controls in all flowering time experiments. Based on SSR results from the study by Thomson et al. (2006), line P9-84 was known to contain an *O. rufipogon* introgression across the *dth1.1a* sub-QTL on chromosome 1, as well as background introgressions on chromosomes 6 and 9; it had been identified as a late flowering line under both SD and LD conditions (Thomson et al. 2006). Lines P4-12 and P15-62 were known to contain *O. rufipogon* DNA across the entire *dth1.1* region and a background introgression on chromosome 2; these lines flowered significantly earlier than the recurrent parent Jefferson as described in Thomson et al. (2006). The inclusion of these controls is necessary when dissecting natural variation in near isogenic lines because background introgressions from a donor may significantly affect phenotypic performance due to G x G and G x E interactions (Yamamoto et al. 2009).

To sub-divide the original *dth1.1* introgression, selfed progenies derived from the pre-SILs were grown in 50-mm-wide x 178-mm-deep plastic pots in the Guterman Greenhouse at Cornell University. DNA was extracted from 6-8 week old seedlings using the Matrix Mill method (Paris and Carter 2000) and marker assisted selection

was performed using SSR markers to identify recombinants containing single- or multiple-candidate genes.

PCR conditions and SSR genotyping across the dth1.1 region: PCR was performed in 15-ml reactions containing 0.2 mM of each SSR or indel primer, 200 mM dNTP mix, 50 mM KCl, 10 mM TRIS-Cl, pH 8.3, 1.5 mM MgCl, 0.01% gelatin, and 1 unit of *Taq* polymerase. The PCR profile was: 94^{o} for 5 min for initial denaturation followed by 35 cycles of 94^{o} for 30 sec, 55^{o} for 30 sec, 72^{o} for 30 sec, and finally 72^{o} for 5 min for final extension. The PCR reaction was performed in a PTC-225 tetrad (MJ Research Watertown, MA) or Mastercycler epgradient (Eppendorf Westbury, NY) thermocycler.

A set of 35 SSRs markers distributed uniformly across the *dth1.1* target region (Fig. 2.2 and Supplementary Table 2.1) was used to determine recombination break points in selfed progenies derived from the pre-SILs. SSR markers were detected using silver stained gels. PCR products were run on 4% denaturing polyacrylamide gels using a manual sequencing gel apparatus followed by silver staining, as described by Panaud et al. (1996). SSR markers were multiplexed three to five times per gel, depending on the length polymorphisms for the Jefferson and *O. rufipogon* alleles at individual loci.

Flowering time gene-specific markers: Primers were designed around indel polymorphisms within flowering time genes on chromosome 1 and on chromosome 6 (*HD1*, *HD3A* and *RFT1*) to distinguish the Jefferson (*tropical japonica*) and *O. rufipogon* alleles (Table 2.1). To identify regions likely to contain indel polymorphisms in the parents of our SILs (cv Jefferson and *O. rufipogon*), we aligned candidate gene sequences from the two fully sequenced rice genomes, cv Nipponbare

(*japonica*) (http://dev.gramene.org/db/cmap/map_set_info?map_set_acc=grjp2008a) and cv 93-11 (*indica*) (http://www.gramene.org/db/cmap/map_set_info ?map_set _acc=bgi2005) and used them as proxies. This strategy was helpful because of the close genetic relationship between Jefferson and Nipponbare, on the one hand, and the relative similarity of *O. rufipogon* (IRGC #105491) and 9311 on the other. The sequences were aligned using the SeqMan program of DNAstar (GeneCodes) and insertions/deletions were identified. Primers flanking the indels were designed using the Primer3 program and tested on the SIL parents. Primers were designed to have a common annealing temperature of 60 ^{o}C and to generate amplicons ~80-240 bp in length. Using BLAST (Altschul et al. 1990), primer sequences were aligned to the sequenced Nipponbare genome to confirm their location and copy number in the rice genome. Indel markers were designed so that the size of the indel polymorphism represented ~10% of the total length of the amplicon to facilitate allele calling on agarose gels. Indel amplicons were size-separated on 2% v/v agarose gels and stained using SYBR green® (http://www.introgen.com).

Background detection using SNP, SSR and RFLP markers: Historical data from 49 RFLP and 103 SSR markers previously mapped onto BC2F2 ancestral materials (Thomson et al. 2003) were used initially to select lines that contained as little *O. rufipogon* DNA in the genetic background as possible. This marker data was supplemented with data from an Illumina Golden Gate SNP assay (http://www.illumina.com) providing information on ~1,300 SNPs well distributed in the rice genome (K. Zhao, Cornell Univ., pers. comm.) (SNP assay developed as part of NSF Plant Genome Award 0606461 to SMc). SNPs included in the Illumina assay were expected to provide enough resolution to estimate the number and size of any spurious introgression(s) remaining in the genetic background of the SILs.

Growth Chamber Evaluation: SILs, parental lines, and early and late flowering controls (including the commercial cultivars, Spring (early) and Madison (late) were evaluated in growth chambers (Conviron Pembina, ND). Plants were evaluated under both LD (14 hrs light) and SD (10 hrs light) conditions, with temperatures of 28 ^{o}C during the light period and 25 ^{o}C during the dark period. Light was provided at an intensity of 100 µmol m^{-2} s^{-1}. A total of 20 plants per line, organized in a randomized complete block design (RCBD), were grown in 50-mm-wide x 178-mm deep plastic pots in growth chambers and all plants were sub-irrigated at a constant water level.

Field Evaluation: SILs, parental lines, pre-SILs, revertants and commercial controls ((cv Jefferson (RA8824), Madison (RA8826), Spring (RA8825), Cocodrie (RA8828) and Wells (RA8827)) were evaluated for flowering time under field conditions during the 2007 and 2008 growing seasons at the USDA-ARS Rice Research Unit in Beaumont, TX and the USDA-ARS Rice Research Unit in Stuttgart, AR. Yield data was based on 4 repetitions in Beaumont, TX during 2007 and in Stuttgart, AR during 2008. Yield data in Stuttgart, AR in 2007 were lost due to extensive bird damage and in Beaumont, TX in 2008 due to hurricane-induced flooding. Nursery-type plots were established by planting 75g/line in six rows in a complete randomized block design (CRBD).

Trait evaluation: Days to flowering (DTF), defined as the number of days from seedling emergence until 50% of the main tillers had spikelets with extruded stigmas, was evaluated in both controlled and field environments, and field-grown plants were additionally evaluated for plant stand (plants/m^2), yield/plant, and yield/plot.

Table 2.1 Flowering time gene allele specific markers were developed by aligning the sequences Nipponbare and 93-11. Identified IN/DEL polymorphisms were used to developed PCR based Candidate gene specific markers to distinguish *O. rufipogon* and Jefferson alleles.

Marker reagent	Forward Primer	Reverse Primer	Amplified Band Size (bp)		TIGR Gene Model	Start (bp)	End (bp)
			Jefferson	*O. rufipogon*			
OsGI	tgaactccatcatgagccacta	acttccagctttgtgcagttg	230	190	LOC_Os01g08700	4,326,087	4,335,288
OsSOC1	tcggcagtggtagagtttga	aaacagaccttgccaccatt	100	70	LOC_Os01g08700	5,466,921	5,469,815
FT-L8	cgacatccttagtgggacaga	ttccttcggtagcatacaacg	160	150	LOC_Os01g10520	5,575,556	5,582,069
FTL	ggctgaaggttttgttttgg	tcatgggttacatgccaattt	190	180	LOC_Os01g11930	6,488,336	6,488,558
OsEMF1	gggggaatttatttcttggt	ggttcgtctacaccagcttc	240	235	LOC_Os01g12890	7,154,582	7,161,187
PNZIP	ttttgaccgaatccatcctt	catcaccttaatggccctgt	90	140	LOC_Os01g17170	9,871,104	9,873,729
RFT1	tggcaagtgagtaaatgaggaa	caaacaccactttttcatgctt	120	131	LOC_Os06g06300	2,925,824	2,927,475
HD3A	tgctcgatcatatcccatctc	ttcggaaagctttctcttttg	90	110	LOC_Os06g06320	2,939,005	2,941,453
HD1	tcgacttgacacccccttac	gcatggctcttgtggaattt	240	205	LOC_Os06g16370	9,335,361	9,337,634

Statistical Analysis: Analysis of variance for all phenotypic characters was performed using the JMP statistical package, version 7.0 for Windows (SAS Institute Inc., Cary, NC). Data from all experiments were normalized by eliminating extreme values and all assumptions of the Least-Square Model (LSM) for controlled environments and Restricted Maximum Likelihood (REML) model for field data were tested to fit phenotypic traits to a linear model and to estimate the variance components. The LSM included the following fixed effects: genotypes (G), Environment (E), replications nested within environments (rep(E)) and genotype by environment interaction (G x E). The REML for flowering time included the following fixed effects: Genotypes (G) and random effects: environment (E), replication nested within environment (rep(E)) and genotype by environment interaction (G x E). The yield model included fixed effects: Genotype (G) and random effects: Environment (E), replication nested within environment (rep(E)) and genotype by environment interaction (G x E).

Multiple means comparisons of all lines for flowering time and yield were done using Dunnet's test with Jefferson as a control ($p<0.05$). Correlation coefficients for plant height, plant stand, yield, panicle length, tiller number, and flowering time were calculated using the same software and the density ellipse command by which Pearson's correlation coefficients were calculated.

Results

Development of candidate gene-containing SILs: After four generations of marker-assisted selection from segregating pre-SIL progeny, eight SILs containing sub-divided *O. rufipogon* introgressions across the *dth1.1* QTL region were selected for further study (Fig. 2.1). Four lines were confirmed to contain introgressions carrying a single candidate gene (*GI, SOC1, EMF1* or *PNZIP*), three carried a pair of

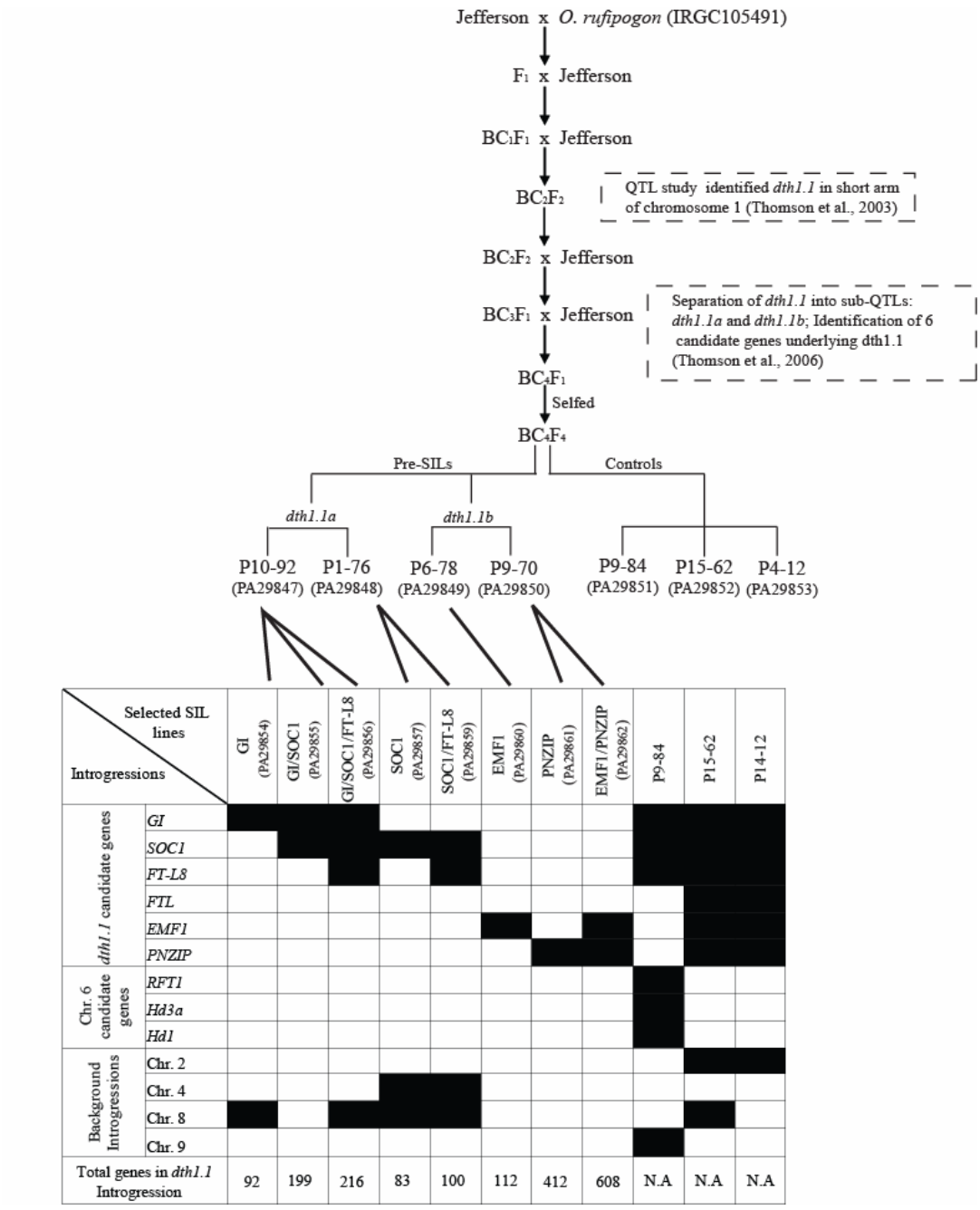
linked candidate genes (*GI/SOC1, SOC1/FT-L8,* or *EMF1/PNZIP*), and one contained a combination of three candidate genes (*GI/SOC1/FT-L8*) (Fig. 2.2). None of the SILs in this study carried an *O. rufipogon* allele at *FTL,* located between *FT-L8* and *EMF1* on chromosome 1.

The previously reported late-flowering control line, P9-84, contained *O. rufipogon* alleles at *OsGI, SOC1* and *FT-L8* in the *dth1.1a* region, as well as a background introgression on the short arm of chromosome 6, containing *HD3A* and *RFT1,* and one on chromosome 9 that contained neither an FT homologue nor a known flowering time gene candidate (Fig. 2.2). Control lines P4-12 and P15-62 both contained a homozygous introgression across the entire *dth1.1* region and were confirmed to carry *O. rufipogon* alleles at all six candidate genes, as well as an introgression on chromosome 2. P15-62 also carried a second background introgression on chromosome 8 (Fig. 2.2). Neither of the introgressions on chromosomes 2 or 8 carried any identifiable flowering time gene candidates.

Four "revertant" pre-SIL controls were selected: P10-92R, P1-76R, P6-78R and P9-70R (Fig. 2.2). These lines were derived by selfing from the backcross populations and contained homozygous Jefferson DNA across the *dth1.1* QTL region (identical to the parental controls), but they also retained a random array of background introgressions inherited from the SIL families (Fig. 2.2). Thus, they allowed us to separate the effect of the target *O. rufipogon* introgressions in the *dth1.1* region from the effect of background introgressions.

To identify background introgressions, the pre-SILs, revertants, eight selected SILs and three controls (P9-84, P15-62 and P14-12) were genotyped with 1,300 evenly distributed SNPs that provided extensive coverage across the rice genome (unpublished data). Our analysis of the SNP data indicated that families associated with the *dth1.1a* QTL carried background introgressions on chromosome 4 and/or 8,

Figure 2.1 Genetic scheme and pedigree showing development of Introgression Lines (SILs) and controls. Four segregating BC_4F_4 pre-SILs, P10-92, P1-76, P6-78 and P9-70, were selected as source materials for SIL development. Summary of candidate flowering time gene(s) and background introgression(s) in each line are indicated in table below.


Jefferson x O. rufipogon (IRGC105491)
F1 x Jefferson
BC1F1 x Jefferson
BC2F2
QTL study identified dth1.1 in short arm of chromosome 1 (Thomson et al., 2003)
BC2F2 x Jefferson
BC3F1 x Jefferson
Separation of dth1.1 into sub-QTLs: dth1.1a and dth1.1b; Identification of 6 candidate genes underlying dth1.1 (Thomson et al., 2006)
BC4F1
Selfed
BC4F4
Pre-SILs
Controls
dth1.1a
dth1.1b
P10-92 (PA29847)
P1-76 (PA29848)
P6-78 (PA29849)
P9-70 (PA29850)
P9-84 (PA29851)
P15-62 (PA29852)
P4-12 (PA29853)
Selected SIL lines
Introgressions
GI (PA29854)
GI/SOC1 (PA29855)
GI/SOC1/FT-L8 (PA29856)
SOC1 (PA29857)
SOC1/FT-L8 (PA29859)
EMF1 (PA29860)
PNZIP (PA29861)
EMF1/PNZIP (PA29862)
P9-84
P15-62
P14-12
dth1.1 candidate genes
GI
SOC1
FT-L8
FTL
EMF1
PNZIP
Chr. 6 candidate genes
RFT1
Hd3a
Hd1
Background Introgressions
Chr. 2
Chr. 4
Chr. 8
Chr. 9
Total genes in dth1.1 Introgression
92
199
216
83
100
112
412
608
N.A
N.A
N.A

while none of the *dth1.1b* QTL families carried any evidence of *O. rufipogon* background introgressions (Figs. 2.1 and 2.2).

Estimates of introgression size: As shown in Fig. 2.2, recombination break points across the *dth1.1* region were mapped using 35 SSR and 86 SNP markers that sub-divided the *dth1.1* region into roughly 700 kb sections (~2.4 cM). By aligning the SSR and SNP positions along the physical map of the sequenced Nipponbare genome (http://www.gramene.org), we were able to estimate the expected size of each introgression in the SILs. Assuming that the physical size of the *O. rufipogon* genome is roughly equivalent to Nipponbare across the *dth1.1* region, we estimate that the introgressions ranged from 0.57 Mb in the SIL containing *GI* to 4.03 Mb in the SIL carrying *EMF/PNZIP*. Similar marker data was used to estimate the size of background introgressions in each of the pre-SILs, which ranged from 260 kb on chromosome 4 to 8.58 Mb on chromosome 6. For ease of reading, individual SILs will hereafter be referred to simply by the name of the candidate gene(s) that they contain, written without italics.

Flowering of SILs under Growth Chambers Conditions: Under SD conditions, the flowering time of the recurrent parent Jefferson was 21.4 days earlier than the *O. rufipogon* donor parent (66.1 vs 87.5 days, respectively; $p<0.001$) (Fig. 2.2). Transgressive variation for early flowering was observed in four of the SILs, each of which contained more than one candidate gene: GI/SOC1 flowered 7.2 days earlier than Jefferson, SOC1/FT-L8 flowered 5.6 days earlier, GI/SOC1/FT-L8 flowered 5.4 days earlier and EMF1/PNZIP flowered 5.9 days earlier. Flowering of the lines harboring single candidate genes was not significantly different than Jefferson under SD (Fig. 2.2).

Figure 2.2 Graphical genotypes of the SILs and controls showing regions of *O. rufipogon* introgression (black rectangle= homozygous; gray=heterozygous) across the *dth1.1a* and *dth1.1b* region of chromosome 1 and the *RFT1/HD3A/HD1* region of chromosome 6. Position of candidate flowering time genes indicated by vertical arrows across top in relation to SSR and SNP markers. Background introgressions indicated to right. Table summarizes days to flowering (DTF) under short day (10 hr.) and long day (14 hr.) and photoperiod sensitivity (PS). Lines that flowered significantly earlier than Jefferson (highlighted in white box) are indicated by light gray rectangles and "*"; late lines indicated by "*L".

Candidate Genes
GI
SOC1
FT-L8
FTL
EMF1
PNZIP
RPT1
HD3A
HD1
dth1.1a
dth1.1b
Chr. 6
Chr. 1
Mbp
SSR
SNP
Background Introgression
Line ID
DTF*
BC4F8 10 hr. Mean
BC4F8 14 hr. Mean
PS*
BC4F8 LD-SD Mean
Pre-SIL
Selected SIL
Controls
Commercial
Revertants
O. rufipogon
Madison
Spring
J/J
R/R
J/R

Under LD, the flowering time of both Jefferson and *O. rufipogon* was delayed, while Jefferson flowered 40 days earlier than *O. rufipogon* ($p<0.0001$). Three SILs showed transgressive variation for earliness under LD; SOC1/FT-L8 flowered 3.4 days earlier and was the only line that flowered earlier than Jefferson under both SD and LD, while both SOC1 and PNZIP flowered 5.7 days earlier (Fig. 2.2). Lines carrying either *GI* or *EMF1*, alone or in combination, were not significantly different than Jefferson. Because combinatorial SILs containing either of these genes were transgressive for early flowering under SD, these results suggest that *O. rufipogon* alleles at *GI* and/or *EMF1* suppress the early flowering response under long days and are thus critical determinants of the differential in flowering time between LD and short days. None of the revertant controls flowered significantly earlier than Jefferson under either SD or LD in growth chamber conditions, indicating that the background introgressions had little effect on flowering time (Fig. 2.2). These results were confirmed in field evaluations where revertant controls flowered similarly to Jefferson with the exception of revertant P10-92R which flowered slightly earlier than Jefferson in Beaumont, TX, but not in Stuttgart, AR.

Photoperiod sensitivity: Photoperiod sensitivity is defined as the difference between days to flowering under LD and SD. Photoperiod sensitivity was detected in all lines tested under growth chamber conditions (Fig. 2.2). Examination of the performance of individual SILs and controls reveals a wide range of responses to photoperiod. The least photoperiod sensitivity was detected in the single candidate gene lines, SOC1 (7.9 days), EMF1 (9.6 days) and PNZIP (5.2 days), and in the commercial, late-flowering control variety, Madison (8.4 days), while high sensitivity to photoperiod was observed in lines EMF1/PNZIP (22.1 days), GI/SOC1 (22.0 days), GI/SOC1/FT-L8 (23.9 days) and the late, pre-SIL P9-84 (22.9 days). Jefferson showed 13.9 days

difference, while donor parent *O. rufipogon* showed the most extreme photoperiod sensitivity, with 32.4 days difference in flowering under SD vs LD conditions (Fig. 2.2). The degree of photoperiod sensitivity was significantly correlated with the number of candidate genes from *O. rufipogon* present in the SILs ($P<0.001$). When two or more candidate genes were present, photoperiod sensitivity of SILs increased by almost 8 days compared with the presence of a single flowering time candidate. This trend was consistent with the fact that SILs containing two or more candidate genes tended to be earlier under SD and later under LD than SILs containing a single candidate gene, which were not significantly different than Jefferson.

Field evaluation of SILs: A summary of the agronomic performance of eight *dth1.1*-derived SILs, two pre-SIL, four revertant controls and five commercial rice varieties, including the Jefferson recurrent parent, is presented in Fig. 2.3. Fitted models for flowering time and yield were highly significant ($p<0.0001$) and explained between 90-95% (R^2 values) of the phenotypic variation for flowering time and ~75% for yield (Table 2.3b and 2.3c).

Under field conditions, the average number of days to flowering for cv Jefferson was 66 days in Beaumont, TX (Lat: 34:13:24 N; Long: 91:31:15 W) and 69 days in Stuttgart, AR (Lat: 30:04:16 N; Long: 94:06:11 W) (Fig. 2.3). This is consistent with the expected difference of ~26 minutes per day in flowering time based on the difference in latitude between the two sites. Among the SILs, SOC1/FT-L8 flowered significantly earlier in both locations while GI/SOC1 flowered 2 days earlier than Jefferson in Beaumont, TX but was similar in Stuttgart, AR. The other five SIL revertants and other controls flowered similarly to Jefferson in both environments with the exception of revertant P10-92R which flowered earlier than Jefferson in Beaumont, TX (Fig.2.3).

Figure 2.3 Bar graph showing days to flowering of SILs and controls in relation to recurrent parent, Jefferson (indicated by block arrow), in Beaumont, TX and Stuttgart, AR. Early, normal and late flowering groups based on mean flowering time across locations.

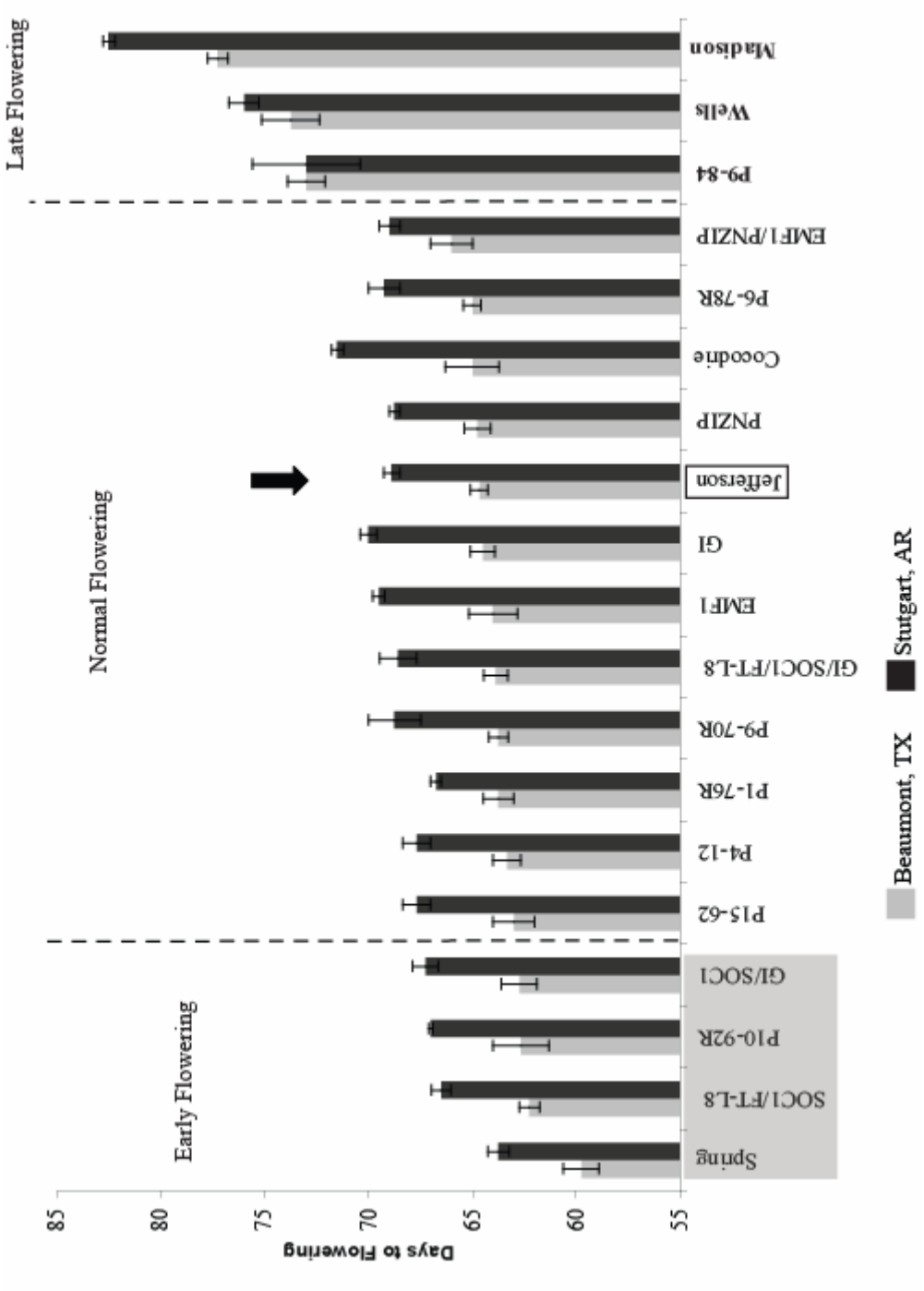

Late Flowering
Normal Flowering
Early Flowering
Madison
Wells
P9-84
EMF1/PNZIP
P6-78R
Cocodrie
PNZIP
Jefferson
GI
EMF1
GI/SOC1/FT-L8
P9-70R
P1-76R
P4-12
P15-62
GI/SOC1
P10-92R
SOC1/FT-L8
Spring
Days to Flowering
85
80
75
70
65
60
55
Beaumont, TX
Stutgart, AR

Table 2.2 Dunnet's multiple means test (P<0.05) of days to flowering in relation to recurrent parent Jefferson (highlighted in white rectangle). Significantly early flowering lines indicated with "*" in gray rectangles; late lines indicated by "*L".

Genotype	**Days to Flowering (DAG)**		
	Beaumont	Sttutgart	Average
Spring	59.8*	63.8*	61.8*
SOC1/FT-L8	62.3*	66.5*	64.4*
P10-92R	62.7*	67.0	64.8*
GI/SOC1	62.7*	67.3	65.0*
P15-62	63.0	67.7	65.4
P4-12	63.3	67.7	65.5
P1-76R	63.8	66.8	65.3
P9-70R	63.8	68.8	66.3
GI/SOC1/FT-L8	63.9	68.6	66.3
EMF1	64.0	69.5	66.8
GI	64.5	70.0	67.3
Jefferson	64.7	68.9	66.8
PNZIP	64.8	68.8	66.8
P6-78R	65.0	69.3	67.2
Cocodrie	65.0	71.5*L	68.3
EMF1/PNZIP	66.0	69.0	67.5
P9-84	73.0*L	75.0*L	74.0*L
Wells	73.8*L	76.0*L	74.9*L
Madison	77.3*L	82.5*L	79.9*L

Pre-SIL P9-84 was consistently later than Jefferson in both environments, and there was a difference of ~10 days between the latest and the earliest Jefferson-derived lines in this experiment (P9-84 and SOC1/FT-L8, respectively).

The most extreme flowering under field conditions was observed in the commercial cultivars. The earliest flowering line was Spring, which flowered two days earlier than the earliest SIL, SOC1/FT-L8. The latest lines were Wells and Madison, which flowered 1-2 weeks later than Jefferson, respectively, and 1-2 days later than the latest line in this study, P9-84 (*O. rufipogon* could not be planted in the field, so is not included in this comparison).

Table 2.3 (A) ANOVA of Days to flowering (DTF) under growth chambers (B) ANOVA of Days to flowering (DTF) under field conditions in Beaumont, TX and Stuttgart, AR. (C) ANOVA for yield under field conditions in Beaumont, TX in 2007 and Stuttgart, AR in 2008.

A

Variance Component	**DTF in Control Environments**			
	SS	% Total	F Value	Prob (F)
δ^2_G	41199.3	26.78	123.3	<.0001
δ^2_{PP}	74699.5	48.55	4249.0	0.0001
$\delta^2_{R(PP)}$	4705.3	3.06	7.9	<.0001
δ^2_{GXPP}	6320.4	4.11	18.9	<.0001

B

Variance Component	**DTF in Field Conditions**			
	Estimate	% Total	F-Value	Prob (F)
δ^2_G	11.88	58.01	92.15	<0.0001
δ^2_L	7.54	36.81	588.42	<0.0001
$\delta^2_{R(L)}$	0.14	0.705	3.97	0.0011
δ^2_{GXL}	-0.08	-0.41	20.74	0.6129

C

Variance Component	**Yield (kg/ha)**			
	Estimate	% Total	F-Value	Prob (F)
δ^2_G	0.18	3.16	2.76	0.0008
δ^2_L	4.42	77.14	77.02	<0.0001
$\delta^2_{R(L)}$	0.25	2.75	2.15	0.0202
δ^2_{GXL}	0.13	2.27	1.55	0.0914

Cocodrie, the highest yielding line, was significantly later than Jefferson in Stuttgart, AR, but flowered similarly to Jefferson in Beaumont, TX (Fig. 2.3).

Agronomic performance of SILs: A preliminary evaluation of yield performance in Beaumont, TX and Stuttgart, AR showed that yield was negatively correlated with flowering time in these SILs ($R = -0.39$, $p < 0.0001$) (Fig. 2.4). These data provide evidence that earliness in these lines does not automatically incur a yield penalty. SILs GI/SOC1, GI/SOC1/FT-L8 and EMF1/PNZIP yielded significantly more than Jefferson as well Madison, Wells and Spring during the 2007 season in Beaumont and were also the highest yielding SILs in Stuttgart in 2008 (Fig. 2.4B). The only commercial cultivar that consistently out-yielded the SILs was Cocodrie, the highest yielding variety in the trials. No differences were observed between the early SILs and the Jefferson recurrent parent for plant height, tiller number, panicle length or panicle number (Supplementary Table 2.2). These results suggest that the *O. rufipogon* alleles at *dth1.1* affect flowering time and yield, but not plant stature, panicle traits or tillering ability.

Discussion

Development of SILs to dissect natural variation: Substitution mapping and phenotypic evaluation of lines carrying non-overlapping *O. rufipogon* introgressions across the *dth1.1* region were previously resolved into at least two separate QTLs (*dth1.1a* and *dth1.1b*) that contributed to earliness (Thomson et al. 2006). In this study, we further dissected the sub-QTLs by generating recombinant SILs, each carrying a single well-defined *O. rufipogon* introgression identified by the presence of one or more flowering time genes.

Figure 2.4 Line graph showing average yield ($kg.ha^{-1}$) of SILs and controls in Beaumont, TX and Stuttgart, AR overlaid on bar graph showing days to flowering in both locations. Early-flowering lines indicated by white rectangles, normal-flowering lines in light gray and late flowering lines in black fill.

Days to Flowering
90.0
85.0
80.0
75.0
70.0
65.0
60.0
55.0
50.0
Yield (kg/ha)
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
Beaumont, TX
Stuttgart, AR
Yield
Cocodrie
GI/SOCI
EMF1/PNZIP
GI/SOCI/FT-L8
P4-12
P6-78R
PNZIP
P10-92R
EMF1
Spring
Wells
P15-62
Madison
P1-76R
SOCI/FT-L8
GI
P9-70R
Jefferson
P9-84

Table 2.4 Summary of yield performance (kg.ha^{-1}) of SILs and controls in relation to recurrent parent Jefferson (highlighted in white rectangle) using Dunnet's multiple means test (P<0.05).

	Yield (kg.ha-1)		
Genotype	**Beaumont**	**Stuttgart**	**Average**
P9-84	3613.6	5486.7	4550.0
Jefferson	3788.2	5681.8	4735.0
GI	4273.0	6572.7	5422.8
SOC1/FT-L8	4643.8	6276.2	5460.0
P1-76R	4300.9	6650.5	5475.7
Madison	4196.8	6828.0	5512.4
P9-70R	5076.1	5595.9	5336.0
P15-62	5216.7	6015.5	5616.1
Wells	4680.4	6674.0	5677.2
Spring	5001.3	6365.3	5683.3
EMF1	5071.7	6314.5	5693.1
P10-92R	4880.7	6528.5	5704.6
PNZIP	5249.0	6212.9	5730.9
P6-78R	5424.8	6153.3	5789.1
P4-12	4921.7	6762.7	5842.2
GI/SOC1/FT-L8	5342.8	6355.0	5848.9*
EMF1/PNZIP	4917.8	6962.8	5940.3*
GI/SOC1	5738.4	6455.5	6097.0*
Cocodrie	6176.1	7334.5	6755.3*

These lines allowed us to observe the effects on flowering time and yield of *O. rufipogon* introgressions containing one or more linked flowering time genes in the Jefferson genetic background. Further dissection of these lines will allow us to fine map the genes underlying the effects on flowering time and yield and to determine how many genes are involved, and whether specific genes are acting pleiotropically on both traits, as in the case of *GHD7* (Xue et al. 2008), or whether genes associated with each trait are simply co-inherited due to linkage.

The individual introgressions in the SILs varied in size between 0.5 and 4.0

Mb and each was predicted to carry between ~90 - 600 genes, based on the annotated Nipponbare gene models. In this study, we used known flowering time genes as sentinels to define the introgressions in the SILs. Using allele-specific gene markers, we confirmed that individual SILs carried the expected donor alleles at each gene.

Our analysis was structured to describe the effect on flowering time of recombinationally defined segments of DNA where each SIL carried a different combination of *O. rufipogon* and Jefferson alleles across the target regions. This work provided an opportunity to ask not only whether *O. rufipogon* alleles in specific sub-regions of the *dth1.1* QTL were capable of generating useful transgressive variation for flowering time, but which specific combinations of Jefferson and *O. rufipogon* alleles across the chromosome 1 QTL and the *HD3A-RFT1*-containing region on chromosome 6 contributed optimally to both flowering time and yield in the southern USA. Prior to the use of molecular markers, it was virtually impossible to analyze whether a particular array of linked genes should be selected intact in a plant breeding program, or whether it should be recombined to achieve maximum advantage. This study offers an example of how recombination within a defined QTL region may be advantageous and how molecular breeding strategies can help identify optimal recombinational profiles.

Promoting fine-scale recombination across the genome is likely to be particularly useful when wild or exotic donors are used as parents in crosses with elite breeding lines. Recombination helps to break up existing linkage blocks related to adaptation and fitness and to generate novel allelic combinations that underlie heterosis and transgressive variation. Because wild and exotic materials are generally highly diverged from elite breeding lines, they provide access to a wider array of allelic variation than do adapted x adapted crosses, but they are often ill-adapted to agronomic environments as evidenced by the fact that interspecific progeny often

succumb to biotic or environmental stress, fail to flower or demonstrate a lack of vigor (fitness) in the new environment. In these cases, genome-wide recombination provides a way of shuffling the genomic deck to generate novel variation and better characterize the value of the exotic materials. Recombination and backcrossing makes it possible to evaluate small regions of donor chromosomes in the context of an adapted genetic background and to reveal the breeding value of the unadapted materials. Targeted, local recombination further helps to mitigate the hitch-hiking effect of deleterious alleles and can fortuitously give rise to valuable new haplotypes that can provide a distinctive advantage in a new genetic background or environment.

Effect of *O. rufipogon* sub-introgressins within *dth1.1:* The most noteworthy result from the dissection of *dth1.1* was the discovery that lines containing introgressions that sub-divided either the *dth1.1a* region containing *GI*, *SOC1* and *FT-L8*, or the *dth1.1b* region containing *EMF1* and *PNZIP*, significantly altered the performance of the lines under SD and LD. Our work strongly suggests that interaction among linked *O. rufipogon* alleles in the introgressed regions contributed positively to early flowering under SD and that *O. rufipogon* alleles contained within the *GI* and *EMF1* introgressions repressed early flowering under LD in the Jefferson background.

Pre-SIL P9-84 contains the same *O. rufipogon* introgression as GI/SOC1/FT-L8, but in addition, contains introgressions on chromosomes 6 (containing *HD3A* and *RFT1* but not *HD1*) and 9 (containing no known flowering time genes). P9-84 was both later flowering and lower-yielding than GI/SOC1/FT-L8. We infer that the late flowering and depressed yields of pre-SIL P9-84 is a consequence of *O. rufipogon* alleles on chromosome 6 and/or 9. P9-84 offers an opportunity to further dissect the gene network governing both yield and flowering time, based on the fact that when *O. rufipogon* alleles are substituted for Jefferson alleles in the introgressed regions on

chromosome 6 and/or 9, the positive transgressive variation for both flowering time and yield conferred by *O. rufipogon* alleles at *dth1.1a* is disrupted.

SOC1/FT-L8 was the only line that flowered earlier than Jefferson in all environments tested, including SD and LD in the growth chamber and field plots in Beaumont, TX and Stuttgart, AR. This line provides a model for studying the genetic control of transgressive variation for flowering time in rice. In this line there are ~100 predicted genes in the ~950 kb introgressed region and it will be interesting to further analyze them to determine which genes/allele(s) from the late flowering *O. rufipogon* parent are necessary to promote early flowering in the Jefferson background and how they function in the context of a genetic network.

Neither of the genes underlying *dth1.1b*, *EMF1* and *PNZIP* have been extensively studied in rice. Both EMF1 and PNZIP SILs flowered similarly to Jefferson, but the combinatorial line EMF1/PNZIP flowered significantly earlier under SD conditions in the growth chamber and out-yielded Jefferson in Beaumont, TX (where days are shorter than in Stuttgart, AR during the summer). Further dissection of the *dth1.1b* region will be needed to better define the genetic factors and functional interactions that promote yield and regulate flowering in these lines.

Role of flowering time genes: The early flowering observed in the SILs cannot be predicted based on the presence or absence of *O. rufipogon* alleles at any one of the flowering time genes investigated here. This argues against the use of a simple "candidate gene" approach to identify the gene(s) responsible for the transgressive phenotype in this study. Instead, we conclude that earliness can result from several different combinations of parental alleles across the *dth1.1a* region, with particular focus on the *O. rufipogon* introgressions marked by *GI/SOC1, SOC1/FTL-8* or *GI/SOC1/FTL-8*. While each introgression contains many genes (making it

impossible to conclude that any one of the flowering time genes is functionally responsible for the phenotype), it is clear that an introgression containing either *SOC1* or *GI* alone does not confer earliness in the Jefferson background, but introgressions containing combinations of these genes, along with *FTL-8* do. Further, we know that an *O. rufipgon* introgression in the *HD3A* region of chromosome 6 can override the effects of the introgressions on chromosome 1, making early lines flower late. In future work, we will examine molecular models that have been developed to predict flowering time in rice to determine whether differences in the expression of the various flowering time genes are predictive of the phenotype, and whether they can provide insight into the genetic basis of transgressive variation for the trait.

Impact of *dth1.1* introgressions on photoperiod sensitivity (PS): Of all the lines and pre-SILs evaluated in this study, SOC1, EMF1 and PNZIP showed the least photoperiod sensitivity in growth chamber conditions. Their degree of photoperiod sensitivity was similar to the late cultivar, Madison. However, all the SILs were earlier than Madison in the field, and they significantly out-yielded Madison. While both lines EMF1 and PNZIP showed low levels of photoperiod sensitivity (9.6 and 5.2 days, respectively), the combinatorial EMF1/PNZIP was one of the most photosensitive of all the lines in this study (22.1 days). Similarly, the single candidate lines, GI, SOC1 and SOC1/FT-L8, had relatively low levels of photoperiod sensitivity (11.9, 7.9 and 16.7 days, respectively), but it was greatly exaggerated in GI/SOC1 (22.0 days) and in GI/SOC1/FT-L8 (23.9 days). This result suggests that gene(s) located within the GI and EMF1 introgressed regions are associated with enhanced photoperiod sensitivity and helps explain why *O. rufipogon* introgressions that include these regions appear to repress early flowering under LD. For example SILs GI/SOC1, GI/SOC1/FT-L8 and EMF1/PNZIP, which are early under SD, are no

longer early under LD, while SOC1/FT-L8 is early under both SD and LD. Using the candidate flowering time genes as reference points, our data fit a linear model that predicts an eight-day increase in photoperiod sensitivity when a SIL contains two or more flowering time genes from *O. rufipogon*, rather than a single gene.

Effect of background introgressions on flowering time and yield: SSR and SNP marker surveys detected and defined background introgressions as summarized in Figure 1. Only the introgression on chromosome 6 carrying *HD3A* and *RTF1* had a significant impact on phenotype and it negatively impacted both flowering time and yield in control line P9-84 in all environments evaluated. Interestingly, no flowering time QTL was identified anywhere on chromosome 6 in the original study (Thomson et al. 2003), despite the fact that this region also contains *HD1 (OsCO),* known to be a major determinant of photoperiod sensitivity and flowering time in rice (Hayama et al. 2003; Takahashi et al. 2009; Yano et al. 2000). The results of this study indicate that *O. rufipogon* alleles in the *HD3A/RTF1*-containing region of chromosome 6 can negate the positive effects of introgressions in the *dth1.1* region and while we did not evaluate the role of an *O. rufipogon* introgression containing *HD1 (OsCO)*(located only 6.4 Mb away from HD3A), this will be investigated in future studies. It is of interest to confirm whether *O. rufipogon* alleles at *HD3A* and *RFT1* are specifically responsible for the late flowering phenotype in P9-84 and, if so, to investigate how they interact with other genes to simultaneously delay flowering and depress yield.

Potential use of *dth1.1* derived SILs: The main rice-producing region of the USA includes the southern states, Arkansas, Louisiana, Texas, Missouri and Mississippi, which account for 60-80% of national production. In this region, early flowering and early maturity are desirable due to the practice of ratoon cropping where two harvests

are obtained from a single planting. Rice is normally planted in this region in March-April and grows vegetatively during the longest days of the year. The crop matures in late August and after the first harvest, plants are left in the field to produce secondary growth, and a ratoon crop is normally harvested in late September. Plant growth late in the season is constrained by low temperatures, so early flowering during the summer is critical to crop productivity, allowing for extended growth duration of the ratoon crop. A general concern about early flowering is the potential reduction in productivity due to a shorter growing season. Thus, it is of interest that most of the *dth1.1* derived lines yielded better than the elite recurrent parent, Jefferson, while flowering at the same time or slightly earlier. This study demonstrates that alleles coming from the late, low-yielding wild donor can enhance the performance of the early, high-yielding USA cultivar, Jefferson.

The transgressive variation captured in the SILs described in this study represents a form of heterosis that does not require the production of F_1 hybrids, as it can be fixed in inbred lines. Transgressive variation is frequently observed in offspring derived from genetically divergent parents and in naturally self-pollinating species where the load of deleterious recessives is low, it can be readily captured in inbred lines. These observations suggest that the transgressive variation for earliness and yield in this study is not due to overdominance. Rather, the underlying genetic mechanism is more likely to be a form of complementary dominance because early flowering was first observed in BC_2F_2 families where the mean family performance was earlier than the early parent, Jefferson (Thomson et al. 2003). We have now fixed the early trait in our SILs and demonstrated that specific combinations of *O. rufipogon* and Jefferson alleles within and between the introgressed regions are predictive of earliness under both LD and SD conditions. We further confirmed the presence of donor alleles at known flowering time genes in each of the target introgressions using

allele specific markers within the genes as "sentinels" during SIL development. This work suggested that *O. rufipogon* does not differ significantly from Jefferson in terms of the identity of genes or the linkage relationships among them in the introgressed regions. Nonetheless, it will be of interest to look carefully at the gene repertoire and gene order in the Jefferson and *O. rufipogon* parents to determine whether novel genetic elements or genome organization may underlie the expression of transgressive variation documented here. Significant differences in genome structure have been documented in both intra- and inter-species comparisons in *Oryza* (Kim et al. 2007; Han and Xue 2003; Huang et al. 2008; Vitte et al. 2007). In addition, it would be of great interest to test our model and determine whether these *O. rufipogon* introgressions confer a similar advantage in different breeding backgrounds with known alleles at the sentinel flowering time genes, and whether F1 hybrid varieties developed using these SILs as one of the parents might further enhance the heterotic expression of earliness and yield.

A long-term objective of this project is to construct introgression lines that can be used as parents in applied plant breeding programs. Because we are working with the cv Jefferson background, it is likely that any lines emerging from this work would be most immediately useful in a program working with *tropical japonica* germplasm and US grain quality. We are also interested in characterizing the gene repertoire found in the *dth1.1* region in the *O. rufipogon* donor used in this study (IRGC 105491), as well as in other wild/weedy accessions of *O. rufipogon,* to understand how the genome is structured across this region and how much variation exists in wild and exotic germplasm. A deeper understanding of how key genes and alleles interact to give rise to transgressive variation for flowering time would allow plant breeders to more effectively manipulate this important reproductive trait in molecular breeding programs.

Acknowledgements

We would like to thank Mrs. Marie Lavallard for the financial support provided for Luis F. Maas' PhD program in the Plant Breeding & Genetics Department at Cornell University through a Frank T. Rhodes Fellowship, the National Science Foundation (Grant DBI 0606461 to SMc) for research funding, Lisa Polewczak for assistance with the field work in the Dale Bumpers National Rice Research Center and the USDA-ARS Rice Research Unit in Beaumont, TX. We also express our gratitude to Michael Thomson and Jeremy Edwards for seeds from the parental pre-SILs provided for this study, and to Michael Gore and Walter de Jong for constructive comments and suggestions during manuscript preparation.

APPENDICES

Appendix Table 2.1 List of indel and SSR markers, chromosomal locations, primer sequences and expected size (bp) in cv Nipponbare.

Marker	Type	Chr	Location (mb)[1]	Forward	Reverse	Expected Size (bp)
OsGI	INDEL	1	4.33	TGAACTCCATCATGAGCCACTA	ACTTCCAGCTTTGTGCAGTTG	230
OsSOC1	INDEL	1	5.47	TCGGCAGTGGTAGAGTTTGA	AAACAGACCTTGCCACCATT	100
FT-L8	INDEL	1	5.58	CGACATCCTTAGTGGGACAGA	TTCCTTCGGTAGCATACAACG	160
FTL	INDEL	1	6.49	GGCTGAAGGTTTTGTTTTGG	TCATGGGTTACATGCCAATTT	190
OsEMF1	INDEL	1	7.15	GGGGGAATTTATTTCTTGGT	GGTTCGTCTACACCAGCTTC	240
PNZIP	INDEL	1	9.87	TTTTGACCGAATCCATCCTT	CATCACCTTAATGGCCCTGT	90
RFT1	INDEL	6	2.93	TGGCAAGTGAGTAAATGAGGAA	CAAACACCACTTTTTCATGCTT	120
HD3A	INDEL	6	2.94	TGCTCGATCATATCCCATCTC	TTCGGAAAGCTTTCTCTTTTG	90
HD1	INDEL	6	9.34	TCGACTTGACACCCCCTTAC	GCATGGCTCTTGTGGAATTT	240
RM462	SSR	1	0.08	ACGGCCCATATAAAAGCCTC	AAGATGGCGGAGTAGCTCAG	243
RM499	SSR	1	0.39	TACCAAACACCAACACTGCG	ACCTGCAGTATCCAAGTGTACG	116
RM3148	SSR	1	0.74	GACTATTGCTCGAACACTTTG	TTGTCTGCTTTGGTATTTGC	166
RM1331	SSR	1	1.67	CACCAGCTTCATGCATGC	AGCACTCAACTGATGCAGTG	139
RM428	SSR	1	2.60	AACAGATGGCATCGTCTTCC	CGCTGCATCCACTACTGTTG	266
RM8105	SSR	1	4.05	TCATTCTCGAAGGCTTACGG	TCAAGCCTAAGCAGAGGATG	117
RM5552	SSR	1	4.21	ATCAGCCCAGAGGGAGTAAC	AGATTCTGGGATCCACGTTG	112
RM6120	SSR	1	4.31	TCAAGAACGAGAAAGCCACC	CCGTGTAGACGACGACGAC	91
RM220	SSR	1	4.42	GGAAGGTAACTGTTTCCAAC	GAAATGCTTCCCACATGTCT	127
RM283	SSR	1	4.88	GTCTACATGTACCCTTGTTGGG	CGGCATGAGAGTCTGTGATG	151

Appendix Table 2.1 (Continue)

RM3233	SSR	1	5.05	GTGGTGAGTAAACAGTGGTGG	GAGAGCAGAGCAGAGGCAAC	117
RM1118	SSR	1	5.56	CCACCAAGCCAAAGAAGAAC	GCTTTCCTTTGTGCTTCTGG	170
RM6277	SSR	1	5.70	TGTCCTTACCCTTGTTTCGC	GTGTGTTCCAACAGTGGTGG	152
RM7466	SSR	1	5.79	CGGTCTGCCTAGCTTGTCTC	ACCGAACACGGAAAAGCC	136
RM620	SSR	1	5.80	GCAACTTCTGGAACTGGATG	GCCTTCTCAGCGCAAAGTC	205
RM621	SSR	1	5.94	CGACAACTTTGAGTGCGAAG	CCATGCATCAACACAACACA	205
RM622	SSR	1	6.15	CAGCCTTGATCGGAAGTAGC	TGCCGTGGTAGATCAGTCTCT	205
RM623	SSR	1	6.22	CATGTGGAAGCCAATCAGAG	ACCAGCGGCACAGTACAAG	205
RM624	SSR	1	6.27	AGATGGTGCAAGCTAAGTTGG	CGCATCAGTTGTTGTCAGTG	205
RM625	SSR	1	6.41	CCTAGCCAGTCCAACTCCTG	GAGTGTCCGACGTGGAGTTC	205
RM626	SSR	1	6.48	TGATGAGGCTCTAGCCGAGT	CATGGACGAAGAAGCAAAGC	205
RM627	SSR	1	6.49	CGTGCGACAGTGGAGTAAAG	AGCTGAGCTGATGGAGAGGA	205
RM628	SSR	1	6.52	AGGCCATAAAGACCACGATG	GATGTTCTCGCTAAGTCTTTCACTC	205
RM629	SSR	1	6.64	GTTCAGGTTTGCAGGTGGAC	TAGCAGCTTGCTTGGATGTG	205
RM490	SSR	1	6.67	ATCTGCACACTGCAAACACC	AGCAAGCAGTGCTTTCAGAG	101
RM1201	SSR	1	7.16	TTACCGCGCCACATATACAC	CGTACGAGCCCTAGTTACCG	186
RM259	SSR	1	7.44	TGGAGTTTGAGAGGAGGG	CTTGTTGCATGGTGCCATGT	162
RM243	SSR	1	7.97	GATCTGCAGACTGCAGTTGC	AGCTGCAACGATGTTGTCC	116
RM8051	SSR	1	9.86	TCTGTTCGATGTTCCCATCG	AAGGAGCGGAAGATCTCCC	153
RM23	SSR	1	10.70	CATTGGAGTGGAGGCTGG	GTCAGGCTTCTGCCATTCTC	145
RM3412	SSR	1	11.57	AAAGCAGGTTTTCCTCCTCC	CCCATGTGCAATGTGTCTTC	211
RM140	SSR	1	12.28	TGCCTCTTCCCTGGCTCCCCTG	GGCATGCCGAATGAAATGCATG	261
RM5365	SSR	1	14.50	TCTGTTCGATGTTCCCATCG	TAAACTCAAACAGGCTGGGC	180
RM5964	SSR	1	17.91	TGATCACCTGCAGGAGCAG	AAGGAGCGGAAGATCTCCC	118

Appendix Table 2.1 (Continue)

RM4949	SSR	2	4.84	CTTTTGGTAATACAGAAGGA	AGACGAAATTTACTGTAGGA	189
RM5862	SSR	2	6.01	TTAGTACCTCATCATAGCTG	CTCTAATCTTCTCTCATTATCA	223
M6734	SSR	6	3.96	TGAGCAGTCTGCAGATGACC	GCTTGGACTTGGAGTCTTGG	197
RM19535	SSR	6	4.97	TAGTATCCCGCTACTCCCTACG	GTCCTGTTACTGCACGACTCC	520
RM276	SSR	6	6.23	CTCAACGTTGACACCTCGTG	TCCTCCATCGAGCAGTATCA	149
RM549	SSR	6	6.98	ACGAACTGATCATATCCGCC	CTGTGGTTGATCCCTGAACC	148
RM19724	SSR	6	8.12	TGCTGGATACTGAAGAGCTGACG	GTGGATTGGTTCCACCACATAGC	345
RM8225	SSR	6	9.31	TGTTGCATATGGTGCTATTTGA	GATACGGCTTCTAGGCCAAA	240
RM19814	SSR	6	10.09	GGGTGAGGAAATGGGAGAGAGG	AAGCAACACACTGGAGAAGTGAGG	230
RM19837	SSR	6	10.97	ACCGAGAAGCAGCTTAAACAACG	GAAGTCGCCTGTATTGTGTGATCC	182
RM19901	SSR	6	12.09	CAGAAATCAGCCAAAGCCTATCG	CCATGACCACCAAAGAAATTCTCC	159
RM544	SSR	8	5.10	TGTGAGCCTGAGCAATAACG	GAAGCGTGTGATATCGCATG	248
RM22558	SSR	8	5.68	GAAGATGTAGAGGTGAACACGAACC	GCCGTTCATCCTACTCCTAATAATGC	180
RM22608	SSR	8	6.66	ACTGTGTCAGGTGCCGTTTAAGG	AATGGCTGGGATTGGGATGG	375
RM22665	SSR	8	8.13	TGCAGGATTGTTGATGAACTCG	CGTCGAATCAAGTTGTGTTGC	333
RM3481	SSR	8	9.13	CTCGTCGCGTTCGTCAAC	CATCTCATCACCTCACGTCG	224
RM22756	SSR	8	10.13	CCGCCTTGTCGTCCATCATATCG	CTCCATCCTCCTCGAGCTCATCC	350
RM2366	SSR	8	11.92	ATTGCCTATATTCATATGGA	GTTATCTGTTACTTCCTTCG	162
RM331	SSR	8	12.29	GAACCAGAGGACAAAAATGC	CATCATACATTTGCAGCCAG	176
RM22870	SSR	8	13.79	CCCGGTAGTAGTGGGTTATGTCC	CTAGTCGCCCTGAGAAGAAGACC	248
RM4595	SSR	8	14.20	AATAGTTGTTGTTTTGGACA	AAATTTAAGTGATTTTGTGC	163
RM22911	SSR	8	15.21	TGAACAAACCGAGAACTGTCTCC	CCAGATCGCATGGAATATATCG	341
RM22951	SSR	8	16.20	GATGAGAGATATTGTGCCGTCTTTCC	CACTGCGTTGCAAGGTTCTCC	286
RM2910	SSR	8	16.81	CAGCTGCTCATATTCATATA	ATAAGGTACTTCATCCGTTA	182

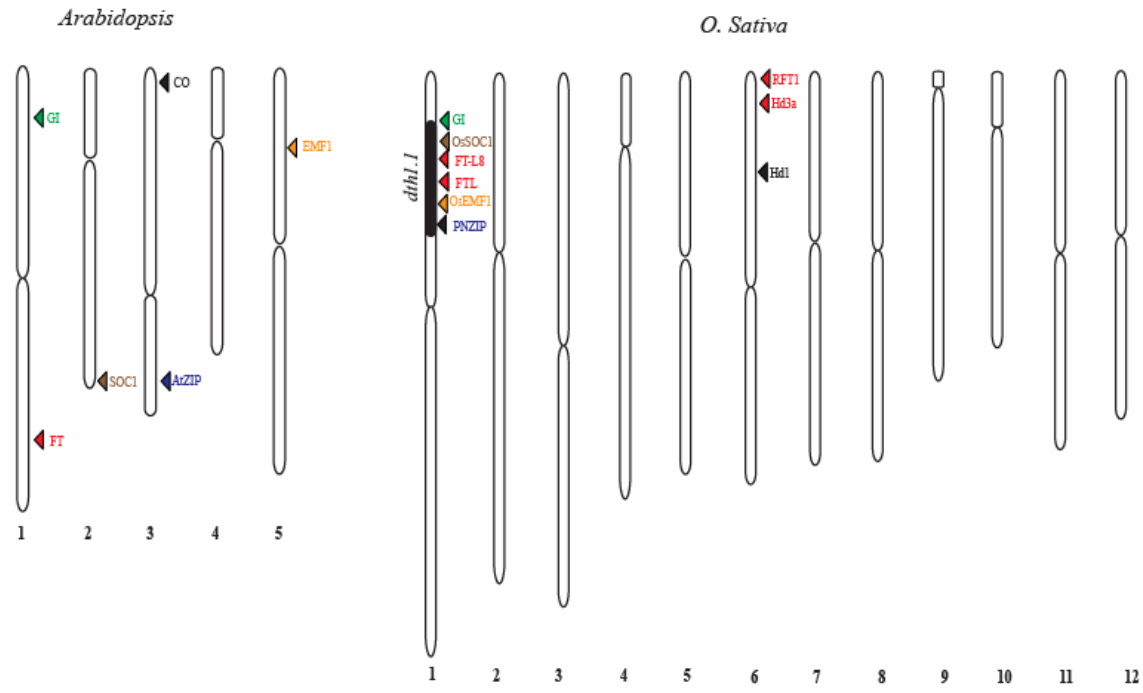
Appendix Table 2.1 (Continue)

RM24122	SSR	9	11.54	GCGGTATGAGTGCGTTTATAGGG	CCTAGTTTACGGATCTGGACATGC	168
RM24181	SSR	9	12.56	ATGCCAGCAAGAAGGCAAGAACG	GACAAGAGTCGCCGAGGCATCC	107
RM12732	SSR	2	6.76	TTACTCCACTTCGCTGATTACTCC	AAATGAGGTACTCCGTCCATCG	468
RM12808	SSR	2	7.75	TACCAGAGTACGGCTGATGC	TCGATTGAGTTGGCTATTGG	420
RM12847	SSR	2	8.53	GGATGTTGTATGTGGTCCCTTGC	CCCACTATGACATAATCCCTCTGC	376
RM12918	SSR	2	9.46	GCAAGTACTATGGGCCTCCAAGC	GGTGGCTGCTACCTCTTATTTAGTCC	279
RM5390	SSR	2	10.81	CTCGACCAAACAGACCAGTAGGG	ATCGCCGCTTAGGAGAATCTGG	107
RM324	SSR	2	11.39	CTGATTCCACACACTTGTGC	GATTCCACGTCAGGATCTTC	175
RM13064	SSR	2	12.41	GGCGTCTACATGACAGACCAATCC	GCTAGAACGAGGAAGAAAGGAAGAGG	362
RM1178	SSR	2	14.01	CAGTGGGCGAGCATAGGAG	ATCCTTTTCTCCCTCTCTCG	112
RM5812	SSR	2	15.89	CGCTGACATCTTGCCCTC	GTAGGACCCACGTGTCATCC	144
RM13213	SSR	2	16.74	GTTTCTCCACCACCGTCAGTCG	CCCTCACTTCACTAGTCCGTAGCC	190
RM335	SSR	4	0.68	GTACACACCCACATCGAGAAG	GCTCTATGCGAGTATCCATGG	104
RM16337	SSR	4	1.72	GCCATCCTCCTCCTAGCAAACC	AAATGTGGTGCTCTGTGAGGTAGC	392
RM16375	SSR	4	2.74	GCCTAACACTCGCTGGAACACC	GACCGGAAAGGGTGAAGAAGG	319
RM16401	SSR	4	3.53	GCAAAGACTTCGCTCCTCTGTACTGC	CCTGTGACAGCGGTGGAGTTCG	459
RM8213	SSR	4	4.43	AGCCCAGTGATACAAAGATG	GCGAGGAGATACCAAGAAAG	177
RM7427	SSR	4	5.67	ACATCGCCGTCCACTCCAC	TCTTCTCCTCCACCCCTACC	135
RM16496	SSR	4	6.49	GATTGGGTGCTAGTGAGCGTAGG	TCTAGAACATGTGTGCGCTTTGC	372
RM5009	SSR	4	7.32	AACCCTAATCTTCACTTCAC	CATTAAACACCTAAATTCCC	165
RM5183	SSR	4	9.25	AATGAGCTAATGTTTCTAAG	AGCTTGAACCTTATATATTG	151
RM3917	SSR	4	11.17	AATGTATTAGGATAAATGCGAAG	GAACGAACGTGAATGAGAAC	262
RM5775	SSR	4	12.42	CACGACCACGACACAAGATG	TGAAGGAACGTTTCTCCTGC	184
RM6467	SSR	6	0.22	GGCAATCTCTCCGAATCTTC	CTAGCTGCTCTGCTCTGCTG	114

Appendix Table 2.1 (Continue)

RM586	SSR	6	1.48	ACCTCGCGTTATTAGGTACCC	GAGATACGCCAACGAGATACC	271
RM5218	SSR	6	2.36	GAAAGCAATTGTTCATTATG	GATAACCCCTGATTTTTTAC	155
RM204	SSR	6	3.17	GTGACTGACTTGGTCATAGGG	GCTAGCCATGCTCTCGTACC	169
RM24231	SSR	9	13.62	CACCACCACCACCCTCTATCTCC	TGAGGCAGGGCCTACTATATCTTGC	293
RM6839	SSR	9	14.57	CTACTGTTGCAGGCTTGCAG	CAGAGGAGGAGATCGAGAGG	104
RM24357	SSR	9	15.56	ACGCTTGCCTTCTCATCGTCTCG	CTGGAACGTCGCCACGTACTGC	177
RM7424	SSR	9	16.58	AGAAGCCCATCTAGCAGCAG	TCAAGCTAGCCACACAGCTG	82
RM257	SSR	9	17.72	CAGTTCCGAGCAAGAGTACTC	GGATCGGACGTGGCATATG	147
RM242	SSR	9	18.81	GGCCAACGTGTGTATGTCTC	TATATGCCAAGACGGATGGG	225
RM160	SSR	9	19.79	AGCTAGCAGCTATAGCTTAGCTGGAG ATCG	TCTCATCGCCATGCGAGGCCTC	131
RM215	SSR	9	21.19	CAAAATGGAGCAGCAAGAGC	TGAGCACCTCCTTCTCTGTAG	148
RM6643	SSR	9	21.71	TGGTGTTATTCCGAGGCTTC	GAGAGAGAGAGGAGATTTTGGG	143
RM2482	SSR	9	22.58	CATGTGCTTTCACAGAAAGT	GGCTCAATGACAACTAAACA	150
RM7586	SSR	9	22.92	GAGTCGTCTCGTCGTAAGCC	TCCTGCCTTCAAACTCGATC	158

A



B

Flowering Time Gene	*Arabidopsis thaliana*			*Oryza sativa*			Function in O. sativa
	Gene Symbol	Chr.	Position	Gene Symbol	Chr.	Position	
GIGANTEA	GI	1	8,061,833	OsGI	1	4,326,162	Circadian clock core
							Regulation of Hd1 expression
CONSTANS	CO	5	24,335,092	Hd1	6	9,335,361	Floral Promotion in short day
							Floral repression in Long day
FLOWERING LOCUS T	FT	1	7,154,582	Hd3a RFT1	6 6	2,939,005 2,925,824	Floral Switch
							Protein Mobile Signal
SUPPRESSOR OF OVEREXPRESSION OF CO1	SOC1/ AGL20	2	18,814,612	OsSOC1/ OsMADS50	3	1,248,228	Floral Promotion
							Repressor of Hd1 under SD
LEU ZIPPER	AtZIP	5	24,977,332	OsZIP (no ortholog)	1	9,863,371	Floral promotion in short day
EMBRYONIC FLOWERING 1	EMF1	5	5,171,185	OsEMF1	1	7,154,582	Floral repressor
							Necessary for vegetative growth

Appendix Figure 2.1 Genomic distribution of homologous flowering time genes in *Arabidopsis* and *O. sativa*.

Appendix Table 2.2 Summary of agronomic trait means in SILs and parental lines.

Line	Height (cm)	Panicle Length (cm)	Panicle Number	Tiller Number
SIL_GI	77.5	16.9	3.1	4.2
SIL_SOC1	75.3	17.7	3.1	3.7
SIL_GI/SOC1	75.1	18.8	3.3	4.2
SIL_SOC1/FT-L8	77.0	17.1	3.3	4.2
SIL_GI/SOC1/FT-L8	77.2	18.1	2.9	3.6
SIL_EMF1	74.6	17.2	3.0	4.3
SIL_PNZIP	75.2	17.3	3.1	3.9
SIL_EMF1/PNZIP	75.8	17.9	3.1	4.4
P9-84	72.4	16.5	2.3	4.6
Jefferson	75.8	18.6	2.1	3.5
O. rufipogon	143.3*	21.5*	1.6*	8.9*

REFERENCES

Altschul SF, Gish W, Miller W, Myer EW, Lipman DJ (1990) Basic local alignment search tool. Mol Biol 215:403-410

Blazquez MA (2000) Flower development pathways. J Cell Sci 113:3547-3548

Cai H, Morishima H (2002) QTL clusters reflect character associations in wild and cultivated rice. Theor Appl Genet 104:1217-1228

Chardon F, Damerval C (2005) Phylogenomic analysis of the PEBP gene family in cereals. J Mol Evol 61:579-590

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M, Yoshimura A (2004) EHD1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of HD1. Genes Dev 18:926-936

Doi K, Yoshimura A, Iwata N (1998) RFLP mapping and QTL analysis of heading date and pollen sterility using backcross populations between *Oryza sativa* L. and *Oryza glaberrima* Steud. Breed Sci 48:395-399

Garner WW, Allard HA (1920) Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. J Agr Res 18:553-606

Han B, Xue Y (2003) Genome-wide intraspecific DNA-sequence variations in rice. Curr Op Plant Biol 6:134-138

Hayama R, Coupland G (2004) The molecular basis of diversity in the photoperiodic flowering responses of *Arabidopsis* and Rice. Plant Physiol 135:677-684

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K (2003) Adaptation of photoperiodic control pathways produces short-day flowering in rice. Nature 422:719-722

Huang X, Lu G, Zhao Q, Liu X, Han B (2008) Genome-wide analysis of transposon insertion polymorphisms reveals intraspecific variation in cultivated rice. Plant Physiol 48:25-40

Izawa T, Oikawa T, Sugiyama N, Tanisaka T, Yano M, Shimamoto K (2002) Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice. Genes Dev 16:2006-2020

Izawa T, Takahashi Y, Yano M (2003) Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and *Arabidopsis*. Curr Op Plant Biol 6:113-120

Kim SL, Lee S, Kim HJ, Nam HG, An G (2007) OsMADS51 is a short-day flowering promoter that functions upstream of EHD1, OsMADS14, and HD3A. Plant Physiol 145:1484-1494

Kohn JR, Leyva N, Dossey R, Sobral B, Morishima H (1997) Quantitative trait locus analysis of trait variation among annual and perennial ecotypes of *Oryza rufipogon*. Int Rice Res Notes 22:4-5

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M (2002) HD3A, a Rice rice ortholog of the *Arabidopsis* FT gene, promotes transition to flowering downstream of HD1 under short-day conditions. Plant Cell Physiol 43:1096-1105

Komiya R, Ikegami A, Tamaki S, Yokoi S, Shimamoto K (2008) HD3A and RFT1 are essential for flowering in rice. Development 135:767-774

Matsubara K, Yamanouchi U, Wang Z-X, Minobe Y, Izawa T, Yano M (2008) EHD2, a Rice ortholog of the maize INDETERMINATE1 gene, promotes flowering by up regulating EHD1. Plant Physiol 148:1425-1435

Panaud O, Chen X, McCouch S (1996) Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). Mol Gen Genet 252:597-607

Paris M, Carter M (2000) Cereal DNA: A rapid high-throughput extraction method for marker assisted selection. Plant Mol Biol Rptr 18:357-360

Park DH, Somers DE, Kim YS, Choy YH, Lim HK, Soh MS, Kim HJ, Kay SA, Nam HG (1999) Control of circadian rhythms and photoperiodic flowering by the

Arabidopsis GIGANTEA Gene. Science 285:1579-1582

Paterson AH, Bowers JE, Feltus FA, Tang H, Lin L, Wang X (2009) Comparative genomics of grasses promises a bountiful harvest. Plant Physiol 149:125-131

Putterill J, Laurie R, Macknight R (2004) It's time to flower: the genetic control of flowering time. BioEssays 26:363-373

Salse J, Piegu B, Cooke R, Delseny M (2002) Synteny between *Arabidopsis thaliana* and rice at the genome level: a tool to identify conservation in the ongoing rice genome sequencing project. Nucleic Acids Res 30:2316-2328

Takahashi Y, Teshima KM, Yokoi S, Innan H, Shimamoto K (2009) Variations in HD1 proteins, HD3A promoters, and EHD1 expression levels contribute to diversity of flowering time in cultivated rice. Proc Natl Acad Sci USA 106:4555-4560

Thomson MJ, Edwards JD, Septiningsih EM, Harrington SE, McCouch SR (2006) Substitution mapping of dth1.1, a flowering-time Quantitative Trait Locus (QTL) associated with transgressive variation in rice, reveals multiple sub-QTL. Genetics 172:2501-2514

Thomson MJ, Tai TH, McClung AM, Lai XH, Hinga ME, Lobos KB, Xu Y, Martinez CP, McCouch SR (2003) Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza*

rufipogon and the *Oryza sativa* cultivar Jefferson. Theor Appl Genet 107:479-493

Tsuji H, Tamaki S, Komiya R, Shimamoto K (2008) Florigen and the photoperiodic control of flowering in rice. Rice 1:25-35

Turck F, Fornara F, Coupland G (2008) Regulation and identity of Florigen: FLOWERING LOCUS T moves center stage. Annl Rev Plant Biol 59:573

Vitte C, Panaud O, Quesneville H (2007) LTR retrotransposons in rice (*Oryza sativa*, L.): recent burst amplifications followed by rapid DNA loss. BMC Genomics 8:218-233

Xiao J, Li J, Grandillo S, Ahn SN, Yuan L, Tanksley SD, McCouch SR (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. Genetics 150:899-909

Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L, Zhou H, Yu S, Xu C, Li X, Zhang Q (2008) Natural variation in GHD7 is an important regulator of heading date and yield potential in rice. Nat Genet 40:761-767

Yamamoto T, Yonemaru J, Yano M (2009) Towards the understanding of complex traits in rice: substantially or superficially? DNA Res 16:141-154

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) HD1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene CONSTANS. Plant Cell 12:2473-2484

CHAPTER 3

ANALYSIS OF EXPRESSION AND NUCLEOTIDE DIVERSITY OF *OsGI*, *HD1*, *HD3A* AND *RFT1* PROVIDE NEW INSIGHTS INTO THE GENETIC BASIS OF TRANSGRESSIVE FLOWERING TIME IN RICE (*ORYZA SATIVA* L.)[2]

Abstract

Many genes in the photoperiodic flowering time pathway have been well-characterized in rice and other plant species using mutants and transgenic lines, but few studies have focused on how recombination among naturally occurring alleles can be used to generate early or late-flowering. In this study, a series of transgressive, early-flowering lines derived from an advanced backcross between *O. sativa,* cv Jefferson (recurrent parent) x *O. rufipogon* (donor) were investigated to understand the relationship between flowering time, DNA sequence variation and expression levels of four flowering time genes (*OsGI*, *HD1*, *HD3A* and *RFT1*). The earliest flowering lines all carried a single introgression corresponding to the QTL, *days to heading 1.1a* (*dth1.1a*) on chromosome 1 from the late-flowering parent, *O. rufipogon.* This introgression contained a number of genes implicated in flowering time control such as *OsGI, OsSOC1* and *FT-L8.* Flowering was delayed in lines carrying a second introgression on chromosome 6 that contained *O. rufipogon* alleles at *HD3A* and *RFT1*. Expression levels of *HD3A* and to a lesser degree *RFT1* were predictive of flowering time, with higher *HD3A* mRNA levels associated with early flowering. These observations provide support for a flowering-time model whereby increases in the expression of *HD3A/RFT1* are caused by trans-acting factors dlocated in the *dth1.1a* QTL region on chromosome 1. In our materials, *O. rufipogon* alleles across the *dth1.1a* QTL combined with Jefferson alleles in the *HD1-HD3A-RTF1* region on chromosome 6 are necessary to drive transgressive early flowering.

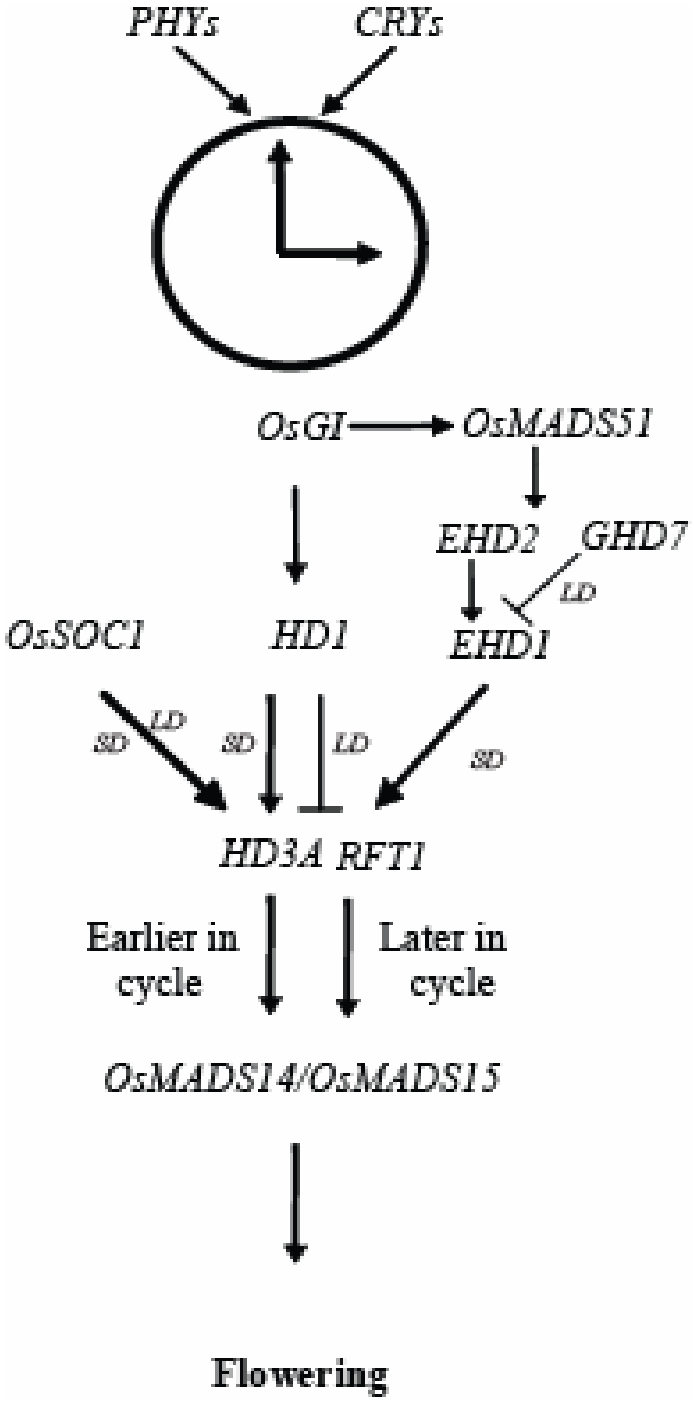
[2]Luis F. Maas, Tesfamichael Kebron, Thomas Brutnell, Chih-Wei Tung and Susan R. McCouch. 2009. Expression and nucleotide diversity of *OsGI*, *HD1*, *HD3A* and *RFT1* contribute to the genetic basis of a QTL for transgressive flowering time in rice (*Oryza sativa* L.). Genetics (to be submitted)

Introduction

Rice (*Oryza sativa* L.) is believed to have been domesticated about 10,000 years ago from the Asian wild rice *Oryza rufipogon,* although the place, time and process by which domestication occurred is still debated (Gao et al. 2008; Khush, 1997; Londo et al. 2006; Vaughan et al. 2007; Ma and Bennetzen, 2004; Vitte et al. 2004). Wild rice species are distributed extensively in tropical and sub-tropical regions of the world (23.5°N and 23.5°S), characterized by warm temperatures, whereas *O. sativa* is cultivated across a broad latitudinal range, between 53°N and 40°S latitudes worldwide (Nanda, 2003). Adaptation of cultivated rice outside of the range of its wild ancestors is largely due to selection for relaxed photoperiod sensitivity and wider variation in flowering time.

Identification of genes involved in the photoperiodic control of flowering time in plants has advanced rapidly during the last decade as a result of numerous forward and reverse genetics studies, coupled with comparative analyses focusing largely on *Arabidopsis* (a facultative long-day plant) and rice (a facultative short-day plant) (Cremer and Coupland, 2003; Hayama et al. 2003; Izawa et al. 2003; Langercrantz, 2009; Tsuji et al. 2008). These studies have highlighted suites of conserved genes and molecular mechanisms that control flowering time in both short and long days. Homologs of the *Arabidopsis* genes *GIGANTEA (GI)*, *CONSTANS (CO)* and *FLOWERING LOCUS T* (*FT)* are present in rice, in the form of *GIGANTEA (OsGI), HEADING DATE 1* (*HD1* or *OsCO*) and *HEADING DATE 3A* (*HD3A* or *OsFT*), respectively (Figure 3.1). However, the function of these core genes is not always conserved between *Arabidopsis* and rice, and help explain the variation in flowering time response between rice and *Arabidopsis* (Hayama et al. 2003; Figure 3.1).

Figure 3.1 Photoperiod pathway in rice. Light signal is perceived by photoreceptors (phytochromes and cryptochromes), that entertains the circadian clock and establish the daily rhythms of gene expression. *GIGANTEA* (*OsGI*) is the immediate output pathway that induces the expression of two alternative flowering induction pathways through *HD1* and *OsMADS51*. *HD1* (the ortholog to Arabidopsis CO) has a dual function by promoting flowering under short-day (SD) and repressing flowering under long-day (LD) by controlling expression of *HD3A* (the ortholog to *Arabidopsis FT*). *HD3A* integrates all signals from all pathways including *OsSOC1*, *HD1* and *EHD1* and its protein represents the mobile signal "florigen" that affects meristem identity genes *OsMADS14/OsMADS15*. *RFT1* is a member of the FT-Like gene family in rice that promotes flowering redundantly than *HD3A*. Up-regulation of *HD3A* and *RFT1* expression associated with early flowering.


PHYs
CRYs
OsGI
OsMADS51
EHD2
GHD7
LD
OsSOC1
HD1
EHD1
SD
LD
SD
LD
SD
HD3A
RFT1
Earlier in cycle
Later in cycle
OsMADS14/OsMADS15
Flowering

As illustrated in Figure 3.1, *OsGI* regulates the expression of *HD1* by integrating signals from light and the circadian clock under both short-day (SD) and long-day (LD) conditions. Over expression of *OsGI,* , delays flowering time under both SD and LD conditions, indicating that the temporal regulation of *OsGI* is an important aspect of flowering time variation (Hayama et al. 2003). *HD1* has a dual function, acting as a promoter of flowering under inductive SD conditions and a repressor of flowering under LD conditions (Hayama et al. 2003; Yano et al. 2000). Under inductive SD conditions, *HD1* expression increases driving the expression of *HD3A* (Hayama et al. 2003). Although the *HD3A* gene is expressed in mature leaves the protein is mobile, signaling the shoot apical meristem to transition from vegetative to the reproductive growth (Tamaki et al. 2007). The rice gene *RICE FLOWERING LOCUS T1* (*RFT1*) encodes a gene with higher similarity to HD3A that FT and has also been shown to act as a floral activator (Izawa, 2007a; Izawa et al. 2002; Kojima et al. 2002; Komiya et al. 2008). *HD3A* and *RFT1* appear to be partially redundant as double knock-out mutations lead to non-flowering phenotypes (Komiya et al. 2008). There are 13 other *FT-like* genes in rice that are also known to be involved in the transition from the vegetative to the reproductive stage, though their specific functions are unknown (Koyima et al. 2008). Activation of *HD3A* and *RFT1* can be achieved independently from *HD1* by either of two genes, *OsSOC1/MADS50* a rice orthologue of *Arabidopsis SUPPRESSOR OF OVEREXPRESSION OF CO1 (SOC1/AGL20)* (Lee et al. 2004; Onouchi et al. 2000; Tagede et al. 2003) or the rice-specific gene, *EARLY HEADING DATE 1* (*EHD1*) (Doi et al. 2004). Overexpression of *OsSOC1* promotes flowering under SD and LD conditions, but its function in the photoperiodic response has not been elucidated (Izawa, 2007b). *EHD1* encodes a B-type response regulator

that is regulated by two upstream genes, *OsMADS51* and *EARLY HEADING DATE 2* (*EHD2*) strictly under SD (Doi et al. 2004; Kim et al. 2007; Matsubara et al. 2008). No homologs of *EHD1*, *OsMADS51* or *EHD2* are found in *Arabidopsis,* indicating that they comprise a unique flowering induction pathway in rice (Fig. 1.) (Doi et al. 2004; Kim et al. 2007; Matsubara et al. 2008).

HD1, *HD3A* and *RFT1* are linked on the short arm of rice chromosome 6 and together play a major role in determining when a particular rice accession will flower (Hayama et al. 2003; Komiya et al. 2008; Yano et al. 2000). Both sequence variation and changes in expression of these flowering time genes have been previously associated with flowering time variation in rice. Functional polymorphism in the coding region of *HD1* and in the promoter of *HD3A* are reported to differentiate between early and late flowering lines (Doi et al. 2004; Hayama et al. 2003; Takahashi et al. 2009; Yano et al. 2000). The second layer of control over flowering time variation is correlated with levels of expression of *HD3A* and *EHD1* (Hayama et al. 2003 and Takahashi et al. 2009).

In a previous study, we identified a flowering time QTL on the short-arm of chromosome 1, *days to heading 1.1a* (*dth1.1a*), where an introgression from the late flowering wild relative *O. rufipogon* (IRGC #105491) conferred transgressive early flowering in the genetic background of the early flowering *tropical japonica* cultivar 'Jefferson' (Thomson et al. 2003; Thomson et al., 2006). Underlying *dth1.1a* is a group of tightly linked genes with high sequence similarity to flowering time genes *OsGI*, *OsSOC1* and *FT-L8* (Thompson et al. 2006). To predict which combination of *O. rufipogon* and cv. Jefferson alleles were responsible for the transgressive flowering time, we developed a series of sub-introgression lines (SILs), each carrying an *O.*

rufipogon introgression at one or more candidate genes underlying *dth1.1a* (Maas et al., 2009). Selected SILs were named according to the candidate gene present in the *O. rufipogon* introgression (Maas et al. 2009). Three SILs having introgressions that contained combinations of two or three flowering time genes, SIL_GI/SOC1, SIL_SOC1/FT-L8 and SIL_GI/SOC1/FT-L8, flowered significantly earlier than the recurrent parent, Jefferson, despite the late-flowering habit of the *O. rufipogon* donor (Maas et al. 2009). The early flowering phenotype of SIL_GI/SOC1/FT-L8 was reversed when an *O. rufipogon* introgression was present in the *HD3A/RFT1* region on chromosome 6. The availability of a set of SILs in a common genetic background provided an opportunity to investigate the relationship between flowering time and expression levels of naturally occurring alleles from *O. rufipogon* and/or *Jefferson* at candidate gene loci on chromosomes 1 and 6.

In this study, we investigate the relationship between sequence variation, gene expression and flowering time in a set of SILs containing different combinations of Jefferson or *O. rufipogon* alleles at *OsGI*, *HD1*, *HD3A* and *RFT1*. We aimed to test the hypothesis that the genetic basis of the transgressive variation for early flowering observed in the interspecific SILs involves an interaction between genes located in the *dth1.1a* QTL region on chromosome 1 and genes in the *HD3A/RFT1* region on chromosome 6 mediated by altered expression levels of key flowering time genes found within the *O. rufipogon* introgressions in these two genomic regions. These studies provide insight into the contributions of allelic variation at several flowering time loci and suggest a mechanism to manipulate flowering time by recombining alleles at tightly linked loci.

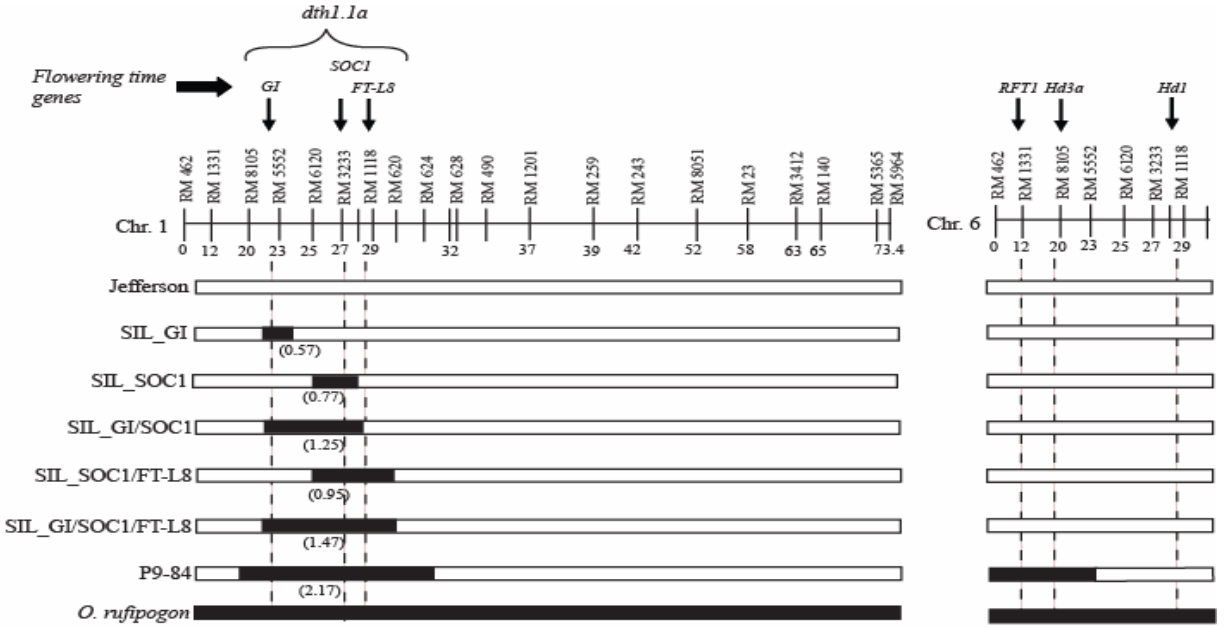
Materials and Methods

Selection of *dth1.1a* derived sub-introgression lines We previously described the selection of near isogenic lines containing well-defined introgressions from the wild relative *O. rufipogon* (IRGC #105491) in the genetic background of the US *tropical japonica* variety Jefferson (Maas et al. 2009) (Fig. 3.2a). Marker assisted selection was employed to genetically dissect the QTL *dth1.1a* into sub-introgression lines (SILs), each carrying a single flowering time candidate gene or combinations of two or three linked candidate genes. This approach allowed us to separate the effects of individual candidate genes and to partition the variation in flowering time observed when defined combinations of Jefferson and *O. rufipogon* alleles were present at each of the candidate loci. After four backcrosses and eight rounds of selfing, five BC_4F_8 SILs were selected for use in this study. The lines are referred to based on the flowering time candidate genes contained within each *O. rufipogon* introgression: SIL_GI, SIL_SOC1, SIL_GI/SOC1, SIL_SOC1/FT-L8 and SIL_GI/SOC1/FT-L8 (Fig. 2a) (Maas et al., 2009). The genetic composition of each line was confirmed using allele-specific markers for the flowering time genes underlying *dth1.1a* QTL, including *OsGI*, *SOC1* and *FT-L8,* as well as the candidate genes on chromosome 6, including *HD1, HD3A* and *RFT1*. Controls for all experiments included the parental lines, *cv* Jefferson (*tropical japonica),* wild rice *O. rufipogon* (IRGC #105491) and commercial cultivars, Madison (late flowering) and Spring (early flowering). A late flowering line, P9-84, was known to contain an *O. rufipogon* introgression across the *dth1.1* sub-QTL containing candidate genes *OSGI, SOC1* and *FT-L8*, one on chromosome 6 carrying the *HD3A* and *RFT1* alleles from *O. rufipogon* and one region on chromosome 9 (containing no known flowering time genes); P9-84 flowered late under SD conditions (Fig. 3.2b).

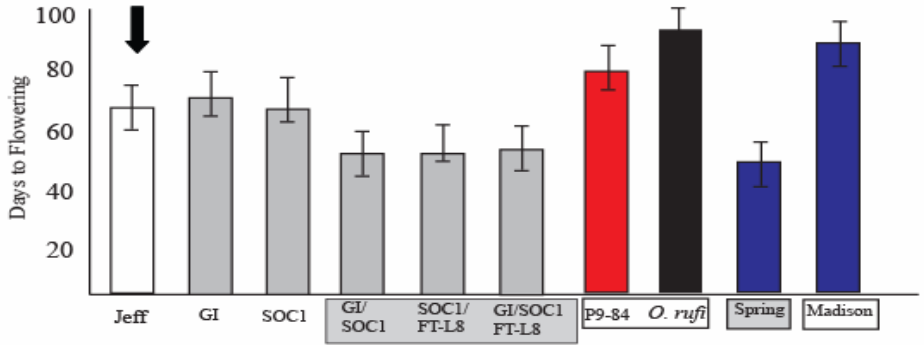
Plant growth conditions and flowering time measurements: SILs, parental lines and early and late flowering commercial controls (Spring and Madison, respectively) were evaluated in growth chambers (Conviron Pembina, ND) in the Guterman greenhouse at Cornell University. Seedlings were grown under day neutral (12L:12D) conditions for 24 days and subsequently exposed to a short-day treatment (10 h light/14 h dark) for 5-9 weeks until heading. The light intensity in the growth chambers was 450 µmol m^{-2} s^{-1}, the temperature was 28 ^{o}C during the light period, 25 ^{o}C during the dark period, and humidity was ~ 70%. Both the temperature and humidity were monitored throughout the growth cycle and data recorded with a HOBO^{TM} RH & TEMP logger (H08-003-02; ONSET Computer Corp.). In the greenhouse environment the temperature averaged 33 ^{o}C day and 28 ^{o}C night, and there were an average of 14 hrs of light under the long-day treatment and 10 hours of light under the short-day treatment. A total of 16-20 plants were evaluated per SIL, organized in a randomized complete block design (RCBD). Plants were grown in 50-mm-wide x 178-mm deep plastic pots in both growth chambers and greenhouse environments and all plants were under-watered at a constant water level.

Sample collection, RNA extraction and quantitative RT-PCR: For gene expression analysis, samples were collected after entraining plants for 10 d under SD conditions in the growth chamber. Mature leaves were simultaneously harvested from five plants in each of three biological replicates for each SIL, frozen in liquid nitrogen and stored in a -80 ^{O}C freezer until RNA could be extracted (Caldana et al., 2007). Analysis of diurnal expression patterns of flowering time genes was carried out at 4 h intervals over a 24 h interval, starting at 07:00 AM when lights were turned on.

Figure 3.2 (A) Graphical genotypes of the SILs and controls showing regions of *O. rufipogon* introgression (black rectangle= homozygous) across the *dth1.1a* region of chr. 1 and the *RFT1/HD3A/HD1* region of chr. 6. Position of candidate flowering time genes indicated by vertical arrows across top in relation to SSR markers. (B) Graphical comparison of days to flowering (DTF) under short day (10 hr.) SILs, parental lines and commercial control Spring and Madison. Lines that flowered significantly earlier than Jefferson (highlighted by block arrow) are indicated by light gray rectangle; late lines indicated by white rectangle.



B. Flowering time comparison of parental lines Jefferson, *O. rufipogon* and dth1.1a derived SILs.



Tissue samples were collected at 0, 4, 8, 12, 16 and 20 h after lights were turned on. To compare expression of genes throughout different growth stages, samples from all lines were also collected at 45, 55, 65 and 75 days after germination (DAG) 4 hours after the lights treatments began. Final gene expression levels were compared among SILs and parental lines using plants at 50 DAG.

We isolated total RNA from leaves using the plant RNeasy minikit (Qiagen, Hilden, Germany). Initial concentrations and quality of extracted RNA were measured using a NanoDrop^{TM} spectrophotometer (Thermo Scientific, Wilmington, DE). Total RNA extraction was straightforward and provided high yield and quality of RNA from leaf tissue (45–75 µg total RNA/100 mg fresh weight). Three micrograms of total RNA were reverse-transcribed using a SuperScript III transcriptase (Invitrogen California, USA) for cDNA synthesis. RNA was diluted using RNAse-free water for quantitative PCR. We carried out quantitative RT-PCR in a total volume of 10 µl containing 1 µl of the reverse-transcribed RNA, 0.25 mM gene-specific primers and 6.25 µl of SYBR Green Master Mix (Applied Biosystems) using the ABI PRISM 7900 HT Real-Time PCR System according to the manufacturer's instructions. The reactions included an initial 5 min of denaturation at 95 ^{o}C, followed by 40 cycles at 95 ^{o}C for 15 s, 55 ^{o}C for 30 s and 72 ^{o}C for 1 min. A dissociation curve was performed to test the specificity of the primers and 10 µl PCR products were separated on a 1% w/v agarose gels.

Primers used included HD1-F and HD1-R for *HD1*, HD3A-F and HD3A-R for *HD3A*, GI-F and GI-R for *OsGI*, FTL8-F and FTL8-R for *FT-L8*, SOC1-F and SOC1-R for *OsSOC1* and Ubiquitin-F and Ubiquitin-R for *Ubquitin3* as the internal control (Supplementary Table 1a). Changes in gene expression were calculated via the $\Delta\Delta_{Ct}$ method (Livak et al*.,* 2001) and the quality of the test was evaluated according to the recommendations of Udvardi et al., (2008).

Selection of reference genes for quantitative RT-PCR: Reference genes were chosen based on published data for rice and other plant species. To identify the most suitable reference genes in rice we compared the performance of five candidate genes for normalization: *ACTIN* (*ACT*), *ACTIN 1* (*ACT1*), *ELONGATION FACTOR 1α* (*EF-1α*), *EXPRESSED PROTEIN* (*EP*) and *UBIQUITIN* (*UBQ*) which are commonly used house-keeping genes in rice (Caldana et al., 2007). Primer sequences for these genes are provided in Supplementary Table 1a. We tested these genes for consistent expression in the parental lines, Jefferson and *O. rufipogon* at different developmental stages as recommended by Caldana et al. (2007) and Udvardi et al. (2008) (Supplementary Fig. 3.1). PCR amplification efficiencies were determined with an ABI PRISM 7900 HT Real-Time PCR system (Supplementary Fig. 3.1). Reaction specificity was evaluated by dissociation curve analysis and by examining amplified products on 1% w/v agarose gels (Supplementary Fig. 3.1).

Of the normalization genes tested, *ACTIN 1* and *EXPRESSED PROTEIN* amplified a secondary band that affected PCR efficiency; thus, they were discarded from further consideration. *UBIQUITIN*, *ACTIN* and *ELONGATION FACTOR 1α* all amplified consistently, representing good candidates for normalization genes. We decided to use *UBQ* as the best normalization gene for expression studies in our population because of its high PCR efficiency (1.95) and specificity (Supplementary Fig. 3.1).

Sequencing of flowering time genes *OsGI*, *HD1, HD3A* and *RFT1:* Sequencing of flowering time genes was accomplished using primers that amplified overlapping gene segments (Supplementary Table 3.1b). The PCR products were amplified using

UltraPFU high fidelity polymerase (Stratagene) and sequenced on an ABI3730 at the Cornell BioResource Center. Each amplicon was sequenced twice, once with the forward primer and a second time with the reverse primer. The accessions sequenced included parental lines Jefferson (*tropical japonica*) and *O. rufipogon* (IRGC #105491), an early and a late SIL, SIL_GI/SOC1 and P9-84, respectively, and the *aus* cultivar Kasalath as a control. Kasalath was incorporated into the sequence analysis because it is a widely used cultivar in flowering time studies and is more similar to *O. rufipogon* (IRGC #105491) than the *tropical japonica* cultivars used in this study. Candidate gene sequences were obtained directly from the Gramene database (http://www.gramene.org/db/cmap/map_set_info?map_set_acc=gt0506 for *cv* Nipponbare (*temperate japonica*) and http://www.gramene.org/db/cmap/map_set_info?map set_acc=bgi2005 for *cv* 93-11 (*indica*)). Sequences were aligned using the Condon Code Aligner program using default parameters (Codon Code Corporation). Identified polymorphisms were then analyzed using the ExPASy translational tool (http://www.expasy.ch/tools/dna.html) to determine the potential impact on protein function.

Gene annotation within *O. rufipogon* introgressions: To determine the number of genes falling within *O. rufipogon* introgressions in each SIL, we employed the GrameneMART module of the Gramene database version 28 (http://www.gramene.org/biomart/martview/) as it allows to specified the query of rice extensive genetic databases. To avoid inflated total counts of predicted genes in these regions, we used a "single occurrence count" and determined how genes were classified based on their annotated biological function using the Gene Ontology (http://www.animalgenome.org/bioinfo/tools/catego/). The use of single occurrences allowed us to count only

once when multiple paths were found between an ancestral term and a child term in a given database.

Statistical Analysis: Analysis of variance for flowering time was performed using the JMP statistical package, version 7.0 for Windows (SAS Institute Inc., Cary, NC). Data from all experiments were normalized by eliminating extreme values and testing all assumptions of the Least-Square Model (LSM) for growth chamber and greenhouse to estimate the variance components. The LSM included the following fixed effects: Genotypes (G), Environment (E), replications nested within Environments (rep(E)) and Genotype by Environment interaction (G x E). Multiple means comparisons of all lines for flowering time and yield were done using the Student T ($p<0.0001$) and Dunnet's test with Jefferson as the control at an error rate of $p<0.05$.

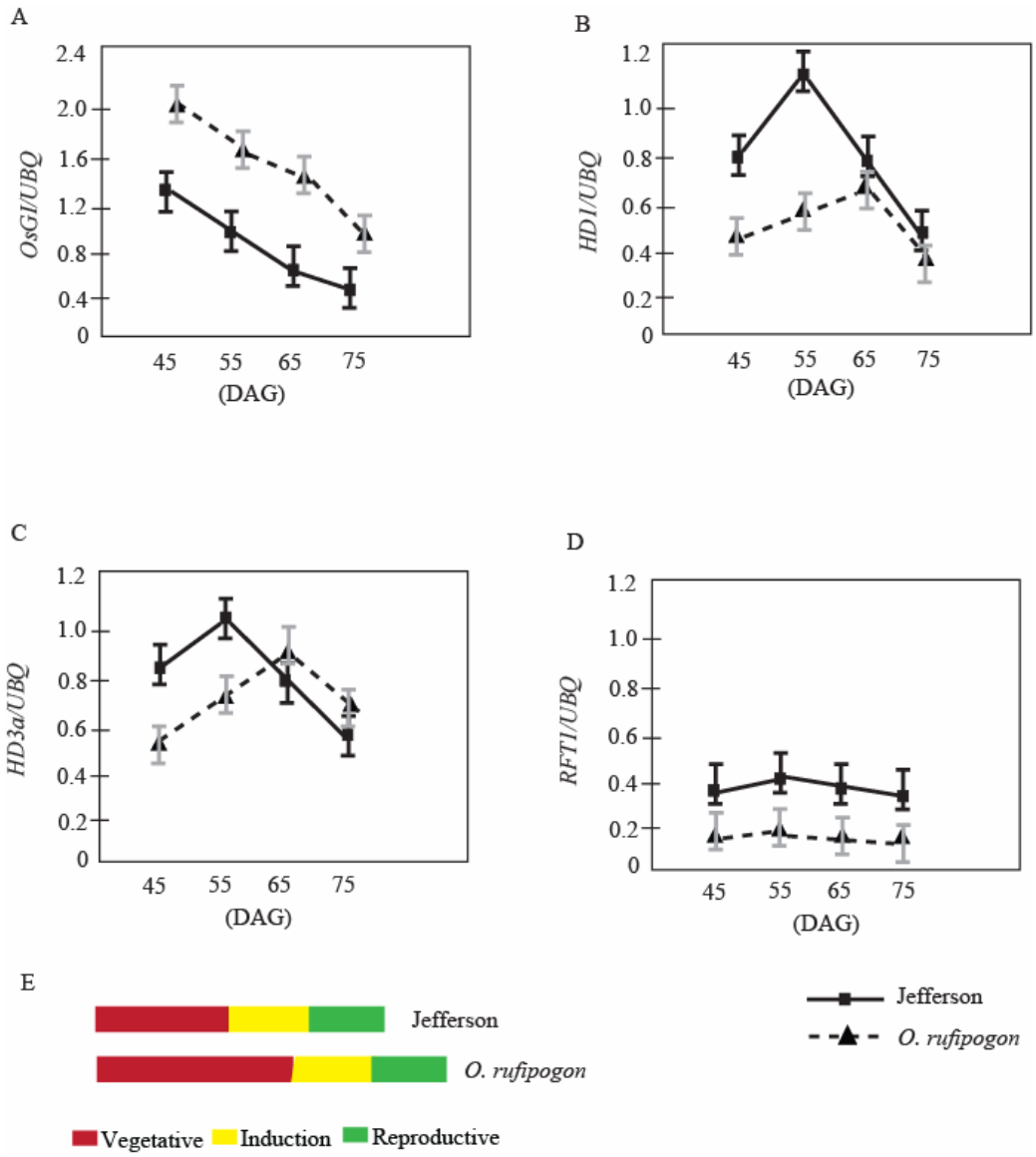
Results

Development of flowering time gene expression platform: As the parental lines displayed vastly different flowering times, it was first necessary to survey gene expression throughout the vegetative stage of growth and at different times of day to standardize the expression analysis process and reduce the sources of variation (Caldana et al. 2007 and Hayama et al. 2003). RT-PCR analysis indicated that the amount and timing of peak expression of *OsGI*, *HD1*, *HD3A* and *RFT1* differed significantly between the two parental lines, Jefferson and *O. rufipogon* (Fig. 3.3a, 3.3b, 3.3c and 3.3d). Levels of *OsGI* mRNA were consistently higher in the late flowering parent, *O. rufipogon* (Fig. 3.3a) while levels of *HD1* (Os*CO*) and *RFT1* were higher in the early flowering cv. Jefferson at all time points analyzed (Fig. 3.3b and 3.3c). For *HD3A,* levels of mRNA peaked around 55 DAG in Jefferson and then

dropped off sharply, while *HD3A* expression in *O. rufipogon* peaked around 65 DAG and dropped off more slowly (Fig. 3.3b and 3.3c). The peak expression of *HD3A* was associated with the transition from vegetative to reproductive stage in both parental lines, occurring approximately 15 days before panicle exertion in each parent under SD conditions. Levels of mRNA measured for *OSGI* were almost twice the levels observed for *HD1* or *HD3A* in both parental lines, while *RFT1* expression levels were very low for both parents and showed almost no change over time, though the difference between the parents was highly significant. Expression of *HD3A* was highly correlated with the expression of *HD1* and the decline in mRNA levels for both *HD1* and *HD3A* occurred rapidly after reaching peak expression Figure 3.3b and 3.3c.

The diurnal expression patterns of *OsGI*, *HD1*, *HD3A* and *RFT1* were investigated by quantifying the relative abundance of mRNA transcripts in leaves of 50 day old plants of Jefferson and *O. rufipogon* grown under SD conditions. Leaves were sampled at 0, 4, 8, 12, 16 and 20 h from the beginning of the light period (Fig. 3.4). *OsGI* expression was higher at all time points in the late flowering *O. rufipogon* (Fig. 3.4a) while expression of *HD1*, *HD3A* and *RFT1* was consistently higher in Jefferson at all time points (Fig. 3.4b, 3.4c and 3.4d). Expression curves for each gene were similar in both parents, with peak expression occurring at the same time of day. *OsGI* and *HD1* peaked at 12 h and 8 hr, respectively, showing circadian control of expression as previously reported by Hayama et al. (2003) (Fig. 3.4a and 3.4b). Expression of *HD3A* and *RFT1* peaked at 4 h after dawn, indicating a conserved induction mechanism for both genes similar to that observed by Komiya et al. (2008) (Fig. 3.4c and 3.4d). Levels of mRNA abundance for *HD3A* and *RFT1* were higher during the light period than the dark period in both parental lines (Fig. 3.4b, 3.4c and 3.4d).

Figure 3.3 Expression analysis of flowering time genes in parental lines Jefferson and *O. rufipogon* at four developmental stages: 45, 55, 65 and 75 Days After Germination (DAG). (A) *OsGI*, (B) *HD1*, (C) *HD3A* and (D) *RFT1*. Relative values of the genes and *UBQ* include standard deviation as vertical bars of five independent experiments.


A
OsGI/UBQ
2.4
2.0
1.6
1.2
0.8
0.4
0
45
55
65
75
(DAG)
B
HD1/UBQ
1.2
1.0
0.8
0.6
0.4
0.2
0
45
55
65
75
(DAG)
C
HD3a/UBQ
1.2
1.0
0.8
0.6
0.4
0.2
0
45
55
65
75
(DAG)
D
RFT1/UBQ
1.2
1.0
0.8
0.6
0.4
0.2
0
45
55
65
75
(DAG)
E
Jefferson
O. rufipogon
Vegetative
Induction
Reproductive
Jefferson
O. rufipogon

Transgressive early flowering of *dth1.1a* SIL associated with higher expression of *HD3A:* Expression levels of *OsGI* were higher in *O. rufipogon* than in any of the other lines in this study (Figure 3.5a). SIL_GI also had significantly higher levels of *OsGI* mRNA than Jefferson, while the other four SILs had expression levels similar to Jefferson, despite the fact that SIL_GI/SOC1 also contained the *O. rufipogon* allele at *OsGI* (Fig. 3.5a).

HD1 expression was dramatically lower in *O. rufipogon* than in any of the other lines, and none of the SILs had *HD1* mRNA levels that were significantly different than Jefferson (Fig. 3.5b). It is noteworthy that none of the SILs carried an *O. rufipogon* allele at the *HD1* gene locus. *HD3A* was the only gene whose expression was significantly correlated with flowering time. *O. rufipogon* and the late-flowering line P9-84 had the lowest levels of mRNA, followed by and SIL_GI, Jefferson and SIL_SOC1, and the three lines showing transgressive variation for earliness, SIL_GI/SOC1, SIL_SOC1/FT-L8 and and SIL_GI/ SOC1/FT-L8 had significantly higher levels of *HD3A* expression (Fig. 3.5c). The expression of *RFT1* was very low in all of the lines at 50 days of age and while the lowest levels of mRNA were observed in *O. rufipogon,* no significant pattern emerged (Fig. 3.5d). It is possible that differences in *RFT1* expression are observed later during development as predicted by Koyima et al. (2008).

These results indicate that increased accumulation of *HD3A* is strongly associated with early flowering and this increase is not directly dependent on enhanced expression of OsGI or *HD1* at 50 days under SD. This suggests that an alternative induction photoperiod pathway in rice, possibly driven by Os*SOC1* and/or *EHD1* mediates the change in *HD3a* expression. Efforts to measure the expression levels of both genes were unsuccessful due to the difficulty in obtaining specific amplification of *OsSOC1* and to the imperceptibly low levels of *EHD1* expression.

Figure 3.4 Diurnal rhythms of (A) *OsGI*, (B) *HD1*, (C) *HD3A* and (D) *RFT1* in parental lines Jefferson and *O. rufipogon* under SD conditions. RNA was extracted from leaf blades of 50 day old plants at 0, 4, 8, 12, 16, and 20 h. Each data point is the average of three independent experiments.

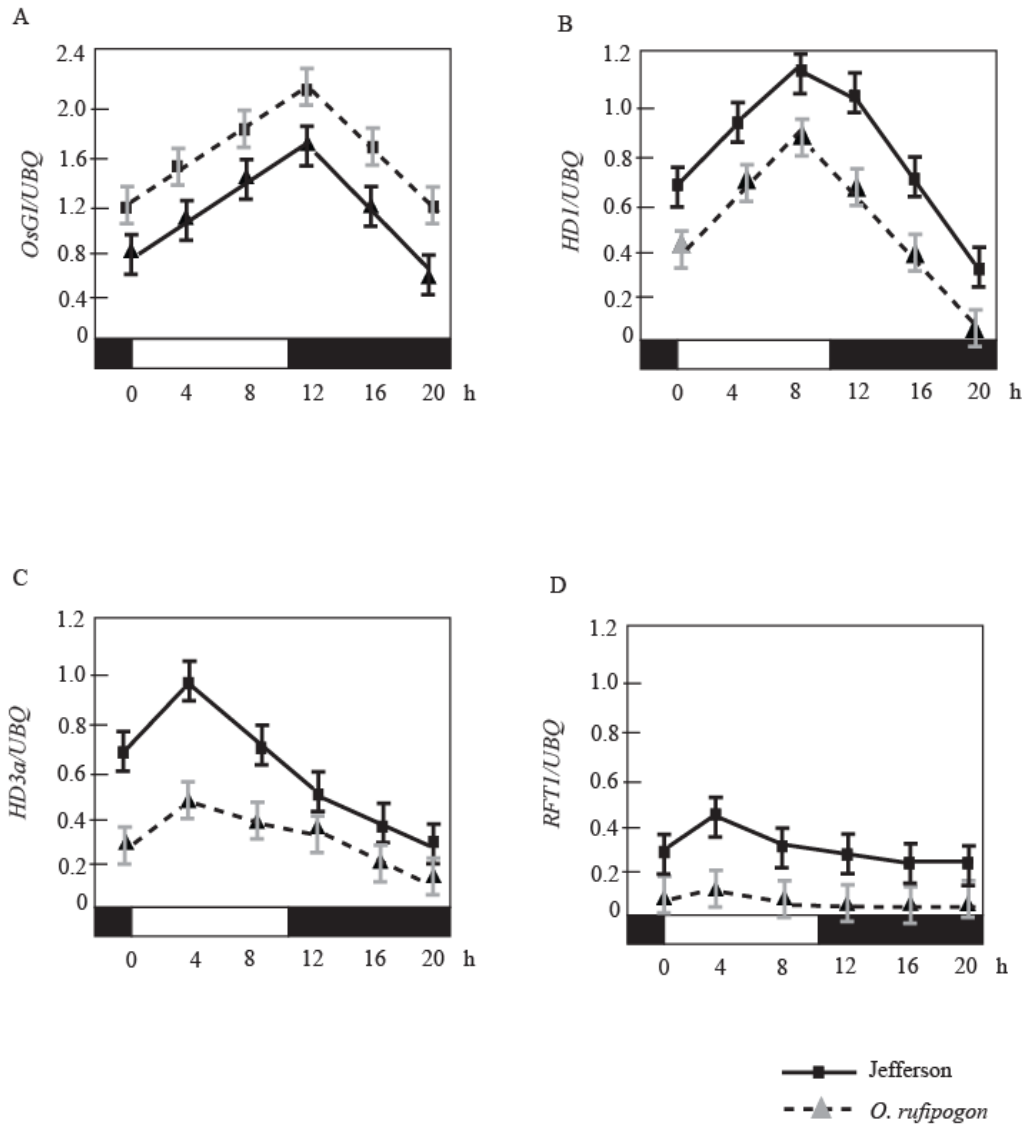
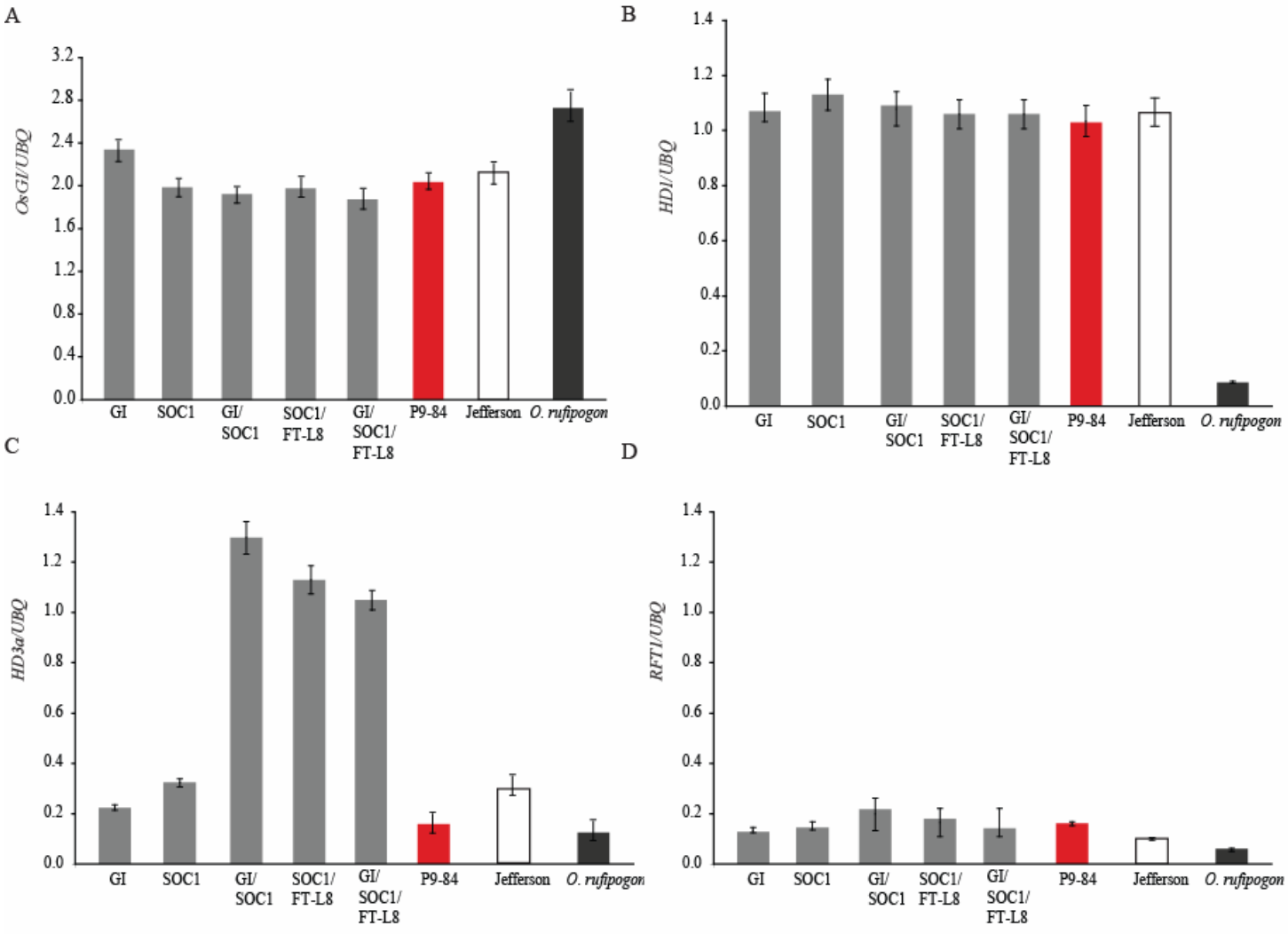

A
OsGI/UBQ
2.4
2.0
1.6
1.2
0.8
0.4
0
0 4 8 12 16 20 h
B
HD1/UBQ
1.2
1.0
0.8
0.6
0.4
0.2
0
0 4 8 12 16 20 h
C
HD3a/UBQ
1.2
1.0
0.8
0.6
0.4
0.2
0
0 4 8 12 16 20 h
D
RFT1/UBQ
1.2
1.0
0.8
0.6
0.4
0.2
0
0 4 8 12 16 20 h
Jefferson
O. rufipogon

Figure 3.5 Expression levels of *OsGI* (A), *HD1* (B), *HD3A* (C) and *RFT1* (D) in parental lines Jefferson and *O. rufipogon*, selected SILs and late line P9-84 under SD conditions (10:14 h day:night). Plants were grown for 50 days and RNA was extracted from leaf blades of five plants harvested 6 h after lights were turned on. Each data point is the average of three independent experiments. Y axis, relative values of transcript levels of flowering time genes in relation to UBQ; standard deviation is represented by vertical bars.


A
OsGI/UBQ
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
GI
SOC1
GI/ SOC1
SOC1/ FT-L8
GI/ SOC1/ FT-L8
P9-84
Jefferson
O. rufipogon
B
Hd1/UBQ
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
GI
SOC1
GI/ SOC1
SOC1/ FT-L8
GI/ SOC1/ FT-L8
P9-84
Jefferson
O. rufipogon
C
Hd3a/UBQ
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
GI
SOC1
GI/ SOC1
SOC1/ FT-L8
GI/ SOC1/ FT-L8
P9-84
Jefferson
O. rufipogon
D
RFT1/UBQ
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
GI
SOC1
GI/ SOC1
SOC1/ FT-L8
GI/ SOC1/ FT-L8
P9-84
Jefferson
O. rufipogon

Nucleotide polymorphisms in flowering time genes: Identification of polymorphisms in diverse lines have led to distinction between early and late flowering phenotypes, that represent a significant layer of variation in cultivated rice. Table 3.1 summarizes the sequence polymorphisms observed between Jefferson and *O. rufipogon* in promoter, exonic and intronic regions of the four flowering time genes, *OsGI, HD1, HD3A* and *RFT1,* in relation to the two reference rice genomes, *cv* Nipponbare (*temperate japonica*) and *cv* 9311 (*indica*). We also sequenced the four genes in the early flowering SIL_GI/SOC1 to confirm the genetic constitution of that line and in the *aus* cultivar, Kasalath, as an additional control. Jefferson was very similar to Nipponbare, with an average of 2.32 indels or point mutations/kb detected across the four loci, while *O. rufipogon* was most closely related to Kasalath, with an average of 1.82 indels or point mutations/kb. We also detected a 9 bp indel in the promoter region of *HD1* that differentiated *O. rufipogon* from Jefferson allele but it was not located near control elements of the promoter region. In addition. both Kasalath and *O. rufipogon* carry a 2 bp deletion in the second exon of *HD1* that introduces an early stop codon disrupting the CCT domain. This polymorphism is characteristic of non-functional types as first reported by Yano et al. (2001) and later confirmed in a diverse group of *temperate japonicas* by Takahashi et al. (2009) (Table 3.1). We also detected a 150 bp insertion in the first exon of the *O. rufipogon HD1* allele that could have an impact on protein functionality (Table 3.1). Together these polymorphisms are likely the reason for the reduced transcript accumulation of *HD1* in *O. rufipogon* via non-sense mediated decay (NMD).

Polymorphisms in the promoter region of *HD3A* have been reported to distinguish between Type A and Type B alleles (Takahashi et al. 2009) where the

Type B promoter carries a 12 bp insertion near the ARR1 binding element and is associated with high *HD3A* mRNA expression levels and early flowering. In our study, the late flowering *O. rufipogon* and late flowering SIL_P9-84 both carry the 12 bp insertion in the *HD3A* promoter but have low levels of *HD3A* transcript. Thus, we conclude that the 12 bp promoter insertion alone is not sufficient for increased levels of of *HD3A* transcripts in transgressive flowering SILs.

Overall, *OsGI* had the lowest level of sequence variation while *RFT1* had the highest (Table 1). Two base substitutions in the 1 kb of *OsGI* promoter differentiated *O. rufipogon* from Jefferson but both were non-synonymous changes and did not explain the higher expression levels of *OsGI* in *O. rufipogon.* Three non-synonymous changes were detected in the coding sequence of *OsGI*, though neither was predicted to cause a loss-of-function in the protein product, and a 255 bp insertion was identified in the eighth intron of the *O. rufipogon* allele. In *RFT1,* we observed a total of 18 substitutions in exons, 14 of which represented non-synonymous amino acid changes, along with 2 deletions of 1 and 11 bp each, which together suggest significant variation in the RFT1 protein. The *O. rufipogon* allele was also distinguished by three indels and 13 substitutions in the promoter, of which 8 were predicted to alter gene expression. Furthermore, 9 indels and 23 substitutions were detected in the introns of *O.rufipogon* allele of *RFT1*compared to Jefferson (Table 1). No causative lesions were associated with any of the tested genes with the exception of *HD1* and *HD3A* in this particular population.

Gene repertoire of individual *O. rufipogon* introgressions: The rice genome sequence was examined to identify gene models that mapped to the introgressed regions in each SIL.

Table 3.1 Sequence variations detected in coding and non-coding regions of flowering time genes *OsGI*, *HD1*, *HD3A* and *RFT1* in parental lines Jefferson and *O. rufipogon*; (a) indicates total number of indel sites, with the number of nucleotides inserted or deleted indicated in parentheses; (b) total number of nucleotide substitutions followed by the number of non-synonymous substitutions in parentheses.

				Promoter			Exon			Intron		
Gene	Species	Variety or Accession	Entire Region (bp)	Length (bp)	Indel[a]	Substitution[b]	Length (bp)	Indel[a]	Substitution[b]	Length (bp)	Indel[a]	Substitution[b]
GIGANTEA	*O. sativa spp. japonica*	Nipponbare	11402	1000	--	--	4639	--	--	6763	--	--
		Jefferson	11402	1000	0	0	4639	0	1 (1)	6763	0	5 (0)
		SIL_GI/SOC1	11147	1000	0	2 (0)	4639	0	5 (3)	5508	1 (255)	11 (0)
	O. sativa spp. indica	Kasalath	11147	1000	0	2 (0)	4639	0	4 (2)	6508	1 (255)	10 (0)
		93-11	11147	1000	0	2 (0)	4639	0	5 (2)	5508	1 (255)	10 (0)
	O. rufipogon	IRGC 105491	11147	1000	0	2 (0)	4639	0	5 (3)	5508	1 (255)	11 (0)
Hd1	*O. sativa spp. japonica*	Nipponbare	5274	1000	--	--	1637	--	--	3637	--	--
		Jefferson	5310	1000	0	1 (0)	1673	1 (36)	3 (0)	3637	0	1 (0)
		SIL_GI/SOC1	5310	1000	0	1 (0)	1673	1 (36)	3 (0)	3637	0	1 (0)
	O. sativa spp. indica	Kasalath	5133	1000	1 (9)	1 (0)	1496	3 (47)	4 (0)	2637	0	6 (1)
		93-11	5133	1000	1 (9)	1 (0)	1496	3 (47)	4 (0)	2637	0	6 (1)
	O. rufipogon	IRGC 105491	5133	1000	1 (9)	1 (0)	1496	4 (203)	4 (0)	2637	0	6 (1)
Hd3a	*O. sativa spp. japonica*	Nipponbare	5499	1000	--	--	847	--	--	4652	--	--
		Jefferson	5499	1000	0	0	847	0	1 (0)	4652	0	5 (0)
		SIL_GI/SOC1	5499	1000	0	0	847	0	1 (0)	4652	0	5 (0)
	O. sativa spp. indica	Kasalath	5480	1000	3 (15)	7 (2)	830	0	5 (1)	4650	2 (14)	16 (2)
		93-11	4480	1000	3 (15)	7 (2)	830	0	5 (1)	3650	2 (14)	16 (2)
	O. rufipogon	IRGC 105491	5480	1000	3 (15)	7 (2)	830	0	5 (1)	4650	2 (14)	16 (2)
RFT1	*O. sativa spp. japonica*	Nipponbare	4652	1000	--	--	866	--	--	3786	--	--
		Jefferson	4652	1000	0	0	866	0	0	3786	0	0
		SIL_GI/SOC1	4652	1000	0	0	866	0	0	3786	0	0
	O. sativa spp. indica	Kasalath	4636	1000	4 (27)	13 (8)	870	2 (12)	18 (14)	3766	9 (42)	23 (17)
		93-11	4636	1000	4 (27)	13 (8)	870	2 (12)	18 (14)	3766	9 (42)	23 (17)
	O. rufipogon	IRGC 105491	4636	1000	3 (23)	13 (8)	870	2 (12)	18 (14)	3766	9 (42)	23 (17)

The number of genes predicted to fall within each introgression and the functional classes of the gene models are documented in Table 2. The number of gene models in each SIL averaged 138, and varied from 83 in the 0.77 Mb introgression containing *OsSOC1* to 216 in the 1.47 Mb introgression containing *GI/SOC1/FT-L8*. About ~60% of the gene models were associated with gene ontology (GO) functional classes involved in biological, cellular and metabolic processes. Genes that would be predicted to have a direct impact on flowering time (i.e., those associated with reproduction, light perception, and flower development) represented less than 5% of the total number, and the flowering time genes used as sentinels in this study represented the strongest candidates found within each introgressed region. The effect of the other genes on flowering time detected in each introgression can not be discarded due to extensive GxG interactions and it suggests that a microarray/Illumina survey of gene expression in the SIL's would be more definitive.

Table 3.2 Functional classification of genes identified in single introgression lines (SILs) based on Gene Ontology (GO)

	Percentage				
GO[a] Classification	GI	SOC1	GI/ SOC1	SOC1/ FT-L8	GI/ SOC1/ FT-L8
Biological process	0.28	0.29	0.29	0.29	0.29
Cellular process	0.14	0.15	0.14	0.16	0.15
Metabolic process	0.16	0.16	0.16	0.16	0.16
Protein metabolic process	0.03	0.03	0.03	0.05	0.04

Table 3.2 (Continue)

Protein modification process	0.01	0.02	0.01	0.03	0.02
Nucleic acid metabolic process	0.07	0.07	0.07	0.06	0.06
Signal transduction	0.01	0.02	0.01	0.03	0.02
Cell communication	0.01	0.02	0.02	0.03	0.02
Response to stress	0.02	0.02	0.02	0.02	0.02
Response to abiotic stimulus	0.02	0.02	0.02	0.01	0.00
Response to endogenous stimulus	0.02	0.02	0.02	0.02	0.02
Transport	0.01	0.02	0.01	0.01	0.01
Biosynthetic process	0.02	0.02	0.02	0.02	0.02
Cellular component organization	0.01	0.00	0.00	0.01	0.01
Transcription	0.05	0.05	0.05	0.04	0.04
DNA metabolic process	0.02	0.00	0.01	0.01	0.01
Precursor metabolites and energy	0.01	0.02	0.01	0.01	0.01
Lipid metabolic process	0.01	0.00	0.00	0.00	0.00
Amino acid and derivative metabolic process	0.02	0.00	0.01	0.00	0.01
Reproduction	0.00	0.00	0.00	0.00	0.00
Translation	0.02	0.01	0.01	0.01	0.01
Flower development	0.00	0.00	0.00	0.00	0.00
Photosynthesis	0.00	0.00	0.00	0.00	0.00
Total	0.93	0.95	0.93	0.96	0.92
Other	0.07	0.05	0.07	0.04	0.08
No. of Genes	92	83	199	100	216

[a](http://www.animalgenome.org/bioinfo/tools/catego/)

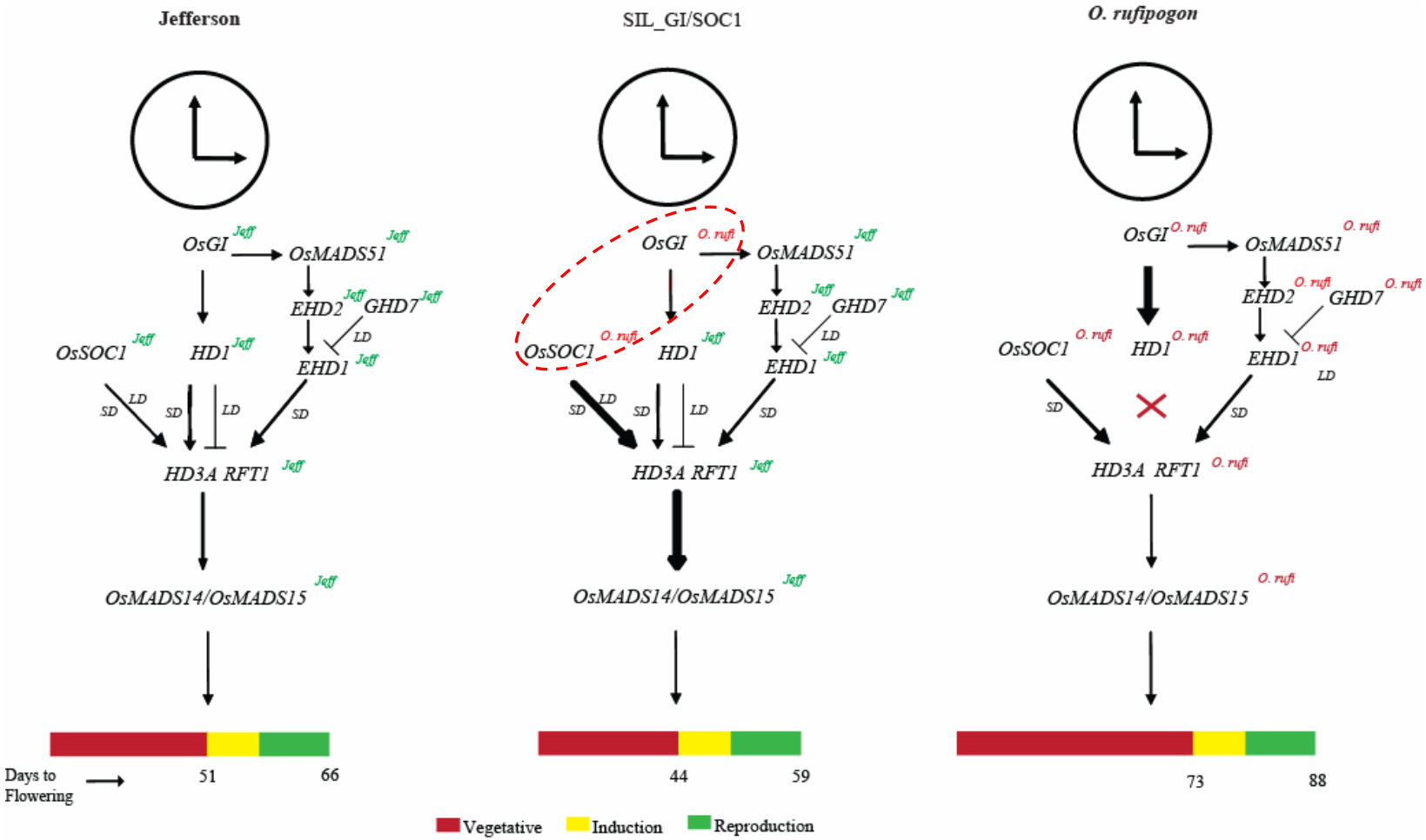
Discussion

Natural genetic variation within a plant species is at the core of breeding for superior genotypes. Whereas much progress has been made in mapping QTLs controlling flowering time, this study seeks to combine expression analysis of candidate flowering time genes on chromosomes 1 and 6 with the genetic dissection of the *dth1.1a* QTL to explore the genetic basis of transgressive variation for early flowering in our collection of SILs. We analyzed gene expression and nucleotide polymorphism for four flowering time genes in a set of sub-introgression lines (SILs) that contained unique combinations of *O. rufipogon* and Jefferson (*tropical japonica*) alleles within the *dth1.1a* QTL region on chromosome 1 (containing *GI, SOC1* and *FT-L8*) and the *HD3A/ RTF1* region on chromosome 6 (Fig. 6). Specifically, we were interested in understanding how alleles from a late flowering, wild relative, *O. rufipogon,* could contribute to transgressive variation for early flowering in the Jefferson recurrent parent.

Our results indicate that differences in flowering time between the early parent, Jefferson, and late flowering parent, *O. rufipogon,* were associated with an earlier transition from the vegetative to the reproductive phase, and are correlated with an increase in the expression of *HD1*, *HD3A* and *RFT1* early in the Jefferson growth cycle. The maximum expression of *HD1* and *HD3A* occurs about 10-20 days sooner in Jefferson than in *O. rufipogon* under SD conditions. The timing of the peak expression of *HD1* and *HD3A* during the plant growth cycle is a useful proxy for

predicting flowering time, as panicle extrusion occurs about 15 days after maximum gene expression (Hayama et al. 2003). The increase in expression of *HD3A* (flowering promoting genes) is correlated with a decrease in the expression of *OsGI*. In previous studies, overexpression of *OsGI* in transgenic rice plants has been shown to delay flowering under both short and long-day photoperiods, acting as a repressor of flowering (Hayama et al. 2003). *OsGI* was the only gene in this study that was expressed at significantly higher levels in *O. rufipogon,* though the elevated expression was not observed in the late line, P9-84. These results suggest that the expression of *OsGI* is determined by interaction with other genes in the flowering time pathway, rather than by the presence or absence of the *O. rufipogon* allele itself at *OsGI*, and that while it may play a central role in delaying flowering in natural populations, late flowering is not necessarily correlated with elevated levels of *OsGI* *RFT1* expression remained at a constant, low level between 45-75 days in both parental lines, suggesting that its role in determining flowering time is not as crucial as that of *HD3A* in this particular population. Changes in the daily patterns of expression, particularly in the moment at which peak expression is observed can be a source of variation in flowering time in mutant rice lines (Izawa et al. 2002; Hayama et al. 2003). However, in this study, the diurnal patterns of expression of *OSGI*, *HD1,* *HD3A* and *RFT1* are conserved in both the cultivated and the wild parent, indicating that changes in the circadian regulation of these genes are not responsible for differences in flowering time. Under SD conditions, the expression of *OsGI* peaks around 12 h after the lights are turned on, and *HD1* peaks at 8 h shortly before the lights are turned off, while *HD3A* and *RFT1* peak about 4 hrs after the lights are turned on, consistent with previous reports (Hayama et al. 2003 and Komiya et al. 2008).

Figure 3.6 Model of photoperiod dependant flowering pathway in parental lines Jefferson and *O. rufipogon* and an early-flowering SIL. Difference in flowering can be attributed to distinctive allelic response of genes *OsGI*, *HD1, HD3A* and *RFT1* that control photoperiodic response in the parental lines. Transgressive variation for early flowering was observed when the Jefferson *HD3A* allele was up-regulated by *O. rufipogon* alleles in the *dth1.1a* introgression on chromosome 1. The replacement of Jefferson alleles at *HD3A/RFT1* with an *O. rufipogon* introgression delayed flowering, making the early plants flower later than Jefferson. *HD3A* is the main gene involved in the transgressive response and early flowering resulted from a combination of *O. rufipogon* alleles in the *dth1.1a* region on chromosome 1 (*GISOC1/FL-L8*) and Jefferson alleles at *HD3A/RFT1* and *HD1* on chromosome 6.


Jefferson
SIL_GI/SOC1
O. rufipogon
OsGI
OsMADS51
EHD2
GHD7
OsSOC1
HD1
EHD1
HD3A RFT1
OsMADS14/OsMADS15
Jeff
O. rufi
LD
SD
Days to Flowering
51
66
44
59
73
88
Vegetative
Induction
Reproduction

Transgressive early flowering in SIL_GI/SOC1, SIL_SOC1/FT-L8 and SIL_GI/SOC1/FT-L8 was associated with significantly elevated *HD3A* expression levels that exceeded those of the early parent, Jefferson, at the flowering induction phase (50 DAG). In contrast, *HD3A* expression levels were not significantly different from Jefferson in SIL_GI and SIL_SOC1, lines that flowered similarly to Jefferson. Interestingly, *HD1* expression was similar in Jefferson and the earliest SILs, indicating that up-regulation of *HD3A* was not dependent on elevated expression of *HD1*, but rather occurred through an alternative pathway such as that involving *OsSOC1* or *EHD1* (Fig. 6). Underlying *dth1.1a* is a candidate gene with sequence similarity to *OsSOC1* and we predict that over-expression of *OsSOC1* up-regulates *HD3A,* leading to early flowering (Lee et al. 2004). We attempted to analyze the expression of *OsSOC1* utilizing the primers and RT-PCR condition established by Komiya et al. (2009), but we were unable to obtain specific amplification of the gene due to the large number of MADS-box genes (>50) present in the rice genome. However, we demonstrated that the presence of the *O. rufipogon* introgression at *OsSOC1* made an essential contribution to transgressive early flowering in SIL_GI/SOC1, SIL_SOC1/FTL-8 and SIL_GI/SOC1/FT-L8 (Maas et al., 2009).

HD3A or *RFT1* function is essential to elicit flowering in rice, as silencing of both genes leads to non-flowering phenotypes (Komiya et al. 2008). *RFT1* appears to play a secondary role compared to *HD3A*, as it is expressed at a much lower level and peak expression is observed ~30 days later in the plant life cycle (Komiya et al. 2008). Results of *RFT1* expression analysis in this study suggest that it does not play a significant role in the induction of flowering at 50 days after germination in the early transgressive lines, as no difference was observed between the earliest SILs and the recurrent parent, Jefferson.

We also investigated the role of these flowering time genes in the delay of flowering. Late flowering in *O. rufipogon* is correlated with higher expression of the flowering repressor gene, *OsGI,* over a longer period of time, as well as a later (delayed) peak of *HD3A* expression (70 DAG) compared to the early parent Jefferson (55 DAG). We infer that the induction of flowering in *O. rufipogon* is regulated primarily through control of *HD3A* expression, where low levels of *HD3A* mRNA and a late peak expression is associated with late flowering. Expression levels of *RFT1* also remained at extremely low levels in *O. rufipogon* for all dates evaluated and this may also contribute to the late flowering of *O. rufipogon*.

The *O. rufipogon* allele of *HD3A* is hypothesized to play a role in delaying flowering, as evidenced by the late flowering behavior of SIL_P9-84. This line flowered significantly later than the recurrent parent Jefferson under all photoperiods, despite the fact that it carried the *dth1.1a* sub-QTL introgression (GI/SOC1/FTL-8) that was associated with transgressive early flowering in SIL_GI/SOC1 /FT-L8. The mechanism by which the *HD3A* allele from *O. rufipogon* retards the onset of the reproductive phase is unknown at this time, but we hypothesize that the *O. rufipogon* allele fails to interact with the induction signal coming from Os*SOC1* or another gene(s) in the *dth1.1a* introgression. Evidence supporting this conclusion is that the transgressive early flowering observed in SIL_GI/SOC1/FT-L8 under short-days was reversed in P9-84 due to an *O. rufipogon* introgression carrying *HD3A*/*RFT1*. Taken together we can propose a model whereby transgressive variation for flowering is based on interaction among *O. rufipogon* alleles at *OsSOC1* (in combination with *OsGI* and/or *FT-L*8 in the *dth1.1a* region on chromosome 1) and Jefferson alleles at *HD3A* and *RTF1* (on chromosome 6). This hypothesis will be more rigorously tested in future experiments. The relationship between expression levels and sequence variation of *HD1* and *HD3A* in relation to flowering time has been investigated in rice

(Doi et al. 2004; Hayama et al. 2003; Takahashi et al. 2009; Tamaki et al. 2007; Yano et al. 2000). Nucleotide diversity in the coding region of *HD1* and the promoter region of *HD3A* are reported to be predictive of variation in flowering time (Doi et al. 2004; Hayama et al. 2003; Takahashi et al. 2009; Tamaki et al. 2007; Yano et al. 2000). Four deletions and a SNP are reported to cause a defect in the CCT domain of the HD1 (OsCO) protein (Takahashi et al. 2009 and Yano et al. 2000) which is known to function as a nuclear localization signal. Lack of a complete CCT domain in *Arabidopsis CO* causes a defective protein (Robson et al. 2001). Sequence analysis in our study confirmed that *O. rufipogon* (IRGC #105491) carries several of the polymorphisms that are associated with the non-functional *HD1* (*OsCO)* allele in rice, while Jefferson carries the same allele as Nipponbare, which is reported to be a functional version. Specifically, the *O. rufipogon* HD1 protein is predicted to have a truncated CCT domain caused by an early stop codon which would subsequently lead to low expression of *HD3A*. In our study, the only line carrying the *O. rufipogon* allele at *HD1* is the *O. rufipogon* parent itself, so we were not able to rigorously test this hypothesis, but it is completely consistent with our data. Further, if we assume that the Jefferson *HD1* allele codes for a functional protein product, our conclusions are consistent with the findings of Takahashi et al. (2009) who showed that flowering time in 64 diverse *temperate japonica* cultivars was highly correlated with a functional *HD1* allele. Takahashi et al. (2009) also identified seven *HD3A* alleles that could be grouped into two types based on their promoter sequences, type A and type B. In general, *HD3A* genes with type B promoters were expressed at significantly higher levels than those with type A promoters. Several potential *cis*-elements in the *HD3A* promoter were analyzed, but no major alterations in potential regulatory sites were identified. This indicates that the association between the type of *HD3A* promoter and *HD3A* expression levels may be caused by other polymorphisms or other tightly

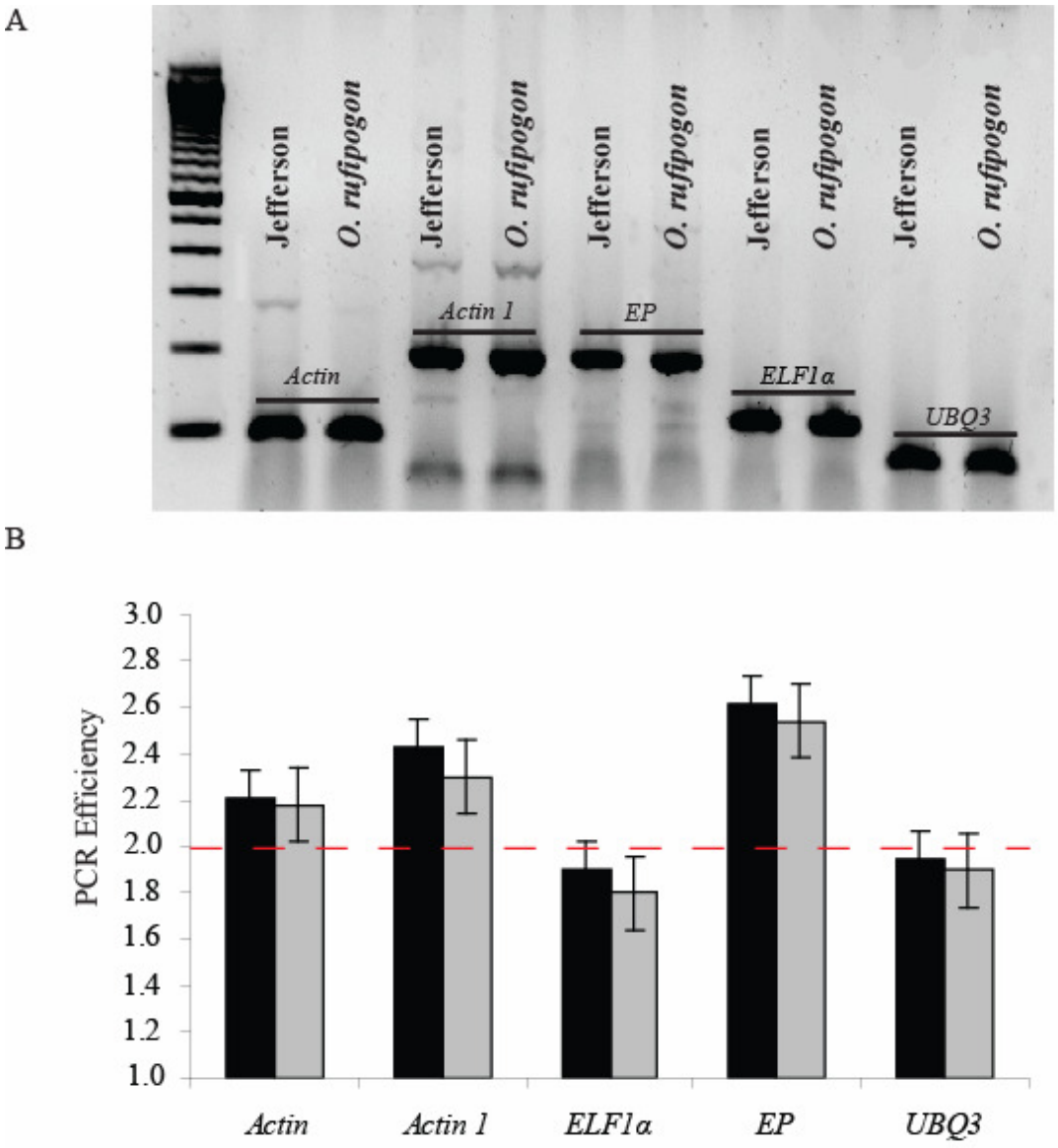
linked loci (Takahashi et al. 2009). Based on polymorphisms detected in the 1 kb promoter region of *HD3A* sequenced in our study, we determined that Jefferson carries a type A promoter and *O. rufipogon* carries a type B promoter, despite the fact that *HD3A* expression levels were significantly lower in *O. rufipogon* than in Jefferson. Low levels of *HD3A* expression were also observed in the late flowering line, P9-84 that carries an *O. rufipogon* introgression across the *HD3A/RFT1* region (type B promoter at *HD3A*) and a Jefferson (functional) allele at *HD1*. This indicates that the known sequence polymorphisms in the *HD3A* promoter region are not sufficient to predict flowering time in rice, even when they occur in combination with a functional allele at *HD1,* suggesting that other factors are required to explain the complex regulation of flowering time under SD in rice. The most important findings in our study are that transgressive flowering time under SD in the early SILs (1) is associated with a four-fold increase in the expression of the flowering promoting gene *HD3A* relative to early parent Jefferson and an 8-fold increase in the expression of *HD1* at 50 DAG relative to late flowering parent *O. rufipogon,* (2) is associated with a simultaneous decrease in the expression of the flowering repressing gene *OsGI* between 45-55 DAG, and that (3) up-regulation of *HD3A* is caused by a pathway other than *HD1* in these materials. Thus it seems that variation in the promoter region of *HD3A* may be the strongest determinate of flowering variation.

To identify the alternative induction pathway of *HD3A* in this population will require further study of *OsSOC1* and *EHD1,* both of which have been shown to induce *HD3A* independently of *HD1*. What still eludes both studies is a clearer understanding of the intrinsic mechanism(s) by which *HD3A* is regulated and how genetics interacts with environmental factors (e.g. light quality, latitude, growths rate) to determine flowering time among rice cultivars. Future studies focused on variation of rice lines carrying the same alleles are needed to deepen our understanding of the control of

flowering time in rice.

Acknowledgements

We would like to thank Mrs. Marie Lavallard and the Department of Plant Breeding and Genetics for financial support for the PhD program of LFM through the Frank T. Rhodes Fellowship. Funding for this research was provided by NSF Plant Genome award #0606461 to SMc. We acknowledge Tesfamichael Kebrom, Dean Ravencroft and Robert Elshire for their assistance in the RT-PCR experiments, to Michael Kovach, Elhan Ersoz and Keyan Zhao for help with the sequence analysis, and to Michael Gore and Walter de Jong for constructive comments and suggestions during manuscript preparation. We are also grateful to Lois Swales for her invaluable assistance in formatting the manuscript.



Appendix Figure 3.1 Normalization gene comparison in parental lines Jefferson and *O. rufipogon*.

Appendix Table 3.1a List of gene specific, normalization and RT-PCR primers.

Primer Name	Forward	Reverse	Product Size
Primers for Allele Identification			
OsGI	TGAACTCCATCATGAGCCACTA	ACTTCCAGCTTTGTGCAGTTG	230
OsSOC1	TCGGCAGTGGTAGAGTTTGA	AAACAGACCTTGCCACCATT	100
FT-L8	CGACATCCTTAGTGGGACAGA	TTCCTTCGGTAGCATACAACG	160
FTL	GGCTGAAGGTTTTGTTTTGG	TCATGGGTTACATGCCAATTT	190
OsEMF1	GGGGGAATTTATTTCTTGGT	GGTTCGTCTACACCAGCTTC	240
PNZIP	TTTTGACCGAATCCATCCTT	CATCACCTTAATGGCCCTGT	90
RFT1	TGGCAAGTGAGTAAATGAGGAA	CAAACACCACTTTTTCATGCTT	120
HD3A	TGCTCGATCATATCCCATCTC	TTCGGAAAGCTTTCTCTTTTG	90
HD1	TCGACTTGACACCCCCTTAC	GCATGGCTCTTGTGGAATTT	240
Normalization Gene Primers			
ACT	CTCCCCCATGCTATCCTTCG	TGAATGAGTAACCACGCTCCG	91
ACT1	ATCCTTGTATGCTAGCGGTCGA	ATCCAACCGGAGGATAGCATG	118
ELF1α	GTCATTGGCCACGTCGACTC	TGTTCATCTCAGCGGCTTCC	118
EP	AGGCTGGTCGAGGAGTCCAT	TTCTCCTCCCTAGCGAACACCT	101
UBQ	AACCAGCTGAGGCCCAAGA	ACGATTGATTTAACCAGTCCATGA	110
Semi-quantitative RT-PCR primers			
OsGI	TGGAGAAAGGTTGTGGATGC	CATAGACGGCACTTCAGCAGAT	209
HD1	GTTTGCAGAGAAGGAAGGGAGCGAG	GGTCGTGCCTCTGCATACGCCTTTCT	401
HD3A	TCAGGGTTTTTTGCAAGATCGATGG	TGGGCTTGGTGCATCTGGGTCTACC	257
UBQ	GACAAGGAGGGAATCCCG	GCATAGCATTTGCGGCA	209
Real time RT-PCR primers			
OsGI	GTGGATGCGCTTTGTGACAT	CGCCTGCAGAAGGATAGGA	69
HD1	AACCAAGATCGGCAGTATGG	GATTGATTGCTCCAGCAGGT	65
HD3A	GCTCACTATCATCATCCAGCATG	CCTTGCTCAGCTATTTAATTGCATAA	67
RFT1	TGGGTTAGCTGACCTAGATTCAAA	GCCAACCACAAGAGGATCGT	63
UBQ	AACCAGCTGAGGCCCAAGA	ACGATTGATTTAACCAGTCCATGA	77

Appendix Table 3.1b List of primers for flowering time gene sequencing.

Primer Name	Forward	Reverse	Product Size
HD1-1	AGGCCTTGCAAACAAACAAA	AATCCTCCCATATGCATTCC	500
HD1-2	CCCTTTGGGCTTAGTTCCAT	TTTTTGTCCTGGGGAGTTTG	523
HD1-3	CAAACTCCCCAGGACAAAAA	CTCAGCGAGGACGGAGGT	594
HD1-4	TCTTGGCTTCTCCTCTCCAA	CATTGACAACGTGGCATGTA	504
HD1-5	TGCCACGTTGTCAATGTTTT	CTGAGATGGAATCGGCAAAT	554
HD1-6	GGGTCTCTGACACCTGCAAT	ACTCCCACTGGATCGATGTT	520
HD1-7	TCCATGGTTCTGATGGGACT	CTGGCCAGTGGACACACTTA	536
HD1-8	CCCTGATTAAGTGTGTCCACTG	AGGGAAGAGGAATGCACTGA	512
HD3A-1	TTAACTAACGGTACGGAAATGG	TCTTGTTGATGATGAAGTGAGGA	507
HD3A-2	ATTGGAGGCAGCAAAAGAGA	AGAGTGAGATGGCCGCTTTA	534
HD3A-3	GGACATGGACATGGACATAGTAA	TGGTCTCTGCACCAACTACG	500
HD3A-4	ACCAGCCTAGGGTCGAGGT	CCCCCAAACAATAGTATGGAAA	560
HD3A-5	TGATTTCCCACTTAAATACATTGC	ATGGAATTCCCATCGAATCA	528
HD3A-6	TCGATGGCACCACACTCTAA	TGTTTAAAAAGTCAACAGCGTCA	504
HD3A-7	TGCCAAGTTTGACCATACCA	CTGGAACAGCACGAACACC	531
HD3A-8	ATGTGCTACGAGAGCCCAAG	AAGAGCACGACTGCATCTCA	519
HD3A-9	AAGTGTGCAATAAGCTGCAAG	CAATCGTACGCGTTTTCTGA	551
HD3A-10	TCAGAAAACGCGTACGATTG	ATTGGCTGAATGCCTGAATC	217
RFT1-1	GAACACACCGCGAATTGTAG	GTGCTGAAACTGACCCCATT	505
RFT1-2	TCTGTCTCGAAATCGCCTCT	TTGTGCAAGCTTCTCTGCAT	505
RFT1-3	CGTGTAGTGTCTTGGGTTGG	CTTCTGGTGGGTCTCTGCAT	587
RFT1-4	GTCGGTGGCAATGACATGAG	ATGGGGTAGGACCAAGGTAA	501
RFT1-5	CATATGTGGCAGTTCCATGA	CTTGGGCTCTCGTAGCACAT	502
RFT1-6	AAATTGATGCAGGGCAAGAG	TGTTTGGCTAGCTTATGAGAAAA	553
RFT1-7	CAAATGAATTATTATTGCAACTGAAAC	TAATTCATCGCCCCCTTTCT	505
RFT1-8	GGGGCGATGAATTATTTTGA	ATCCCATGAATCGACTGCTC	327
OsGI_1	GCGTTGATTCTAGCTGGATTTT	GGTGGAAGCTTTCTCGTTTTTA	709
OsGI_2	AGAAAGCTTCCACCTTTTTCCT	CCCATCAATCCACTTCTCATTT	833
OsGI_3	CAGCTTCAAATGAGAAGTGGAT	AAAGAGATGAAGGACATGAACG	602
OsGI_4	GATAGAAATGGTTTCCCGTTCA	TAATACCCAGAGGTGCAGCAA	472
OsGI_5	GGATCACAGACATATTGCTTGC	AGTTGTTGGAGGCTTCAATTCT	538
OsGI_6	AAAATACGCAGCTGGTGGAG	GGCATAAGTTGTGGGTGCTT	947
OsGI_7	GGGAAGCACCCACAACTTAT	CTAGTGGCTCATGATGGAGTTC	766
OsGI_8	TGAACTCCATCATGAGCCACTA	ACTTCCAGCTTTGTGCAGTTG	858
OsGI_9	CCTTCAACTGCACAAAGCTG	GCTTCCAAGATGCCAAGTATTC	841
OsGI_10	TCGTAGAATACTTGGCATCTTGG	GCAGAACGATAGCAGCTGAAG	871
OsGI_11	GCTTCAGCTGCTATCGTTCTG	GGAGTTTATTTGTCCGCTGTTC	853
OsGI_12	CATCTTGAACAGCGGACAAATA	TATGCGATATTCCGTCGAAAC	750

REFERENCES

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T (2005) *FD*, a *bZIP* Protein Mediating Signals from the Floral Pathway Integrator FT at the Shoot Apex. Science 309: 1052

Andersen CH, Jensen CS, Petersen K (2004) Similar genetic switch systems integrate the floral inductive pathways in dicots and monocots. Trends Plant Sci 9: 105-107

Araki T, Kobayashi Y, Kaya H, Iwabuchi M (1998) The flowering-time gene FT and regulation of flowering in *Arabidopsis*. Journal of Plant Research 111: 277-281

Araki T, Komeda Y (1993) Analysis of the Role of the Late-Flowering Locus, *Gl*, in the Flowering of *Arabidopsis thaliana*. Plant Journal 3: 231-239

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, et al. (2007) Genome-Wide Patterns of Nucleotide Polymorphism in Domesticated Rice. PLoS Genet 3(9): e163

Caldana C, Scheible WR, Mueller-Roeber B, Ruzicic S (2007) A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors. Plant Methods 3:7

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I Giakountis A, Farrona S, Gissot L, Turnbull C , et al (2007) FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of *Arabidopsis*. Science 316: 1030-1033

Cremer F, Coupland G (2003) Distinct photoperiodic responses are conferred by the same genetic pathways in *Arabidopsis* and in rice. Trends Plant Sci 8:405-407

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M , Yoshimura A (2004) *EHD1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT-like* gene expression independently of *HD1*. Genes & Dev 18:926-936

Fowler S, Lee K, Onouchi H, Samach A, Richardson K, Morris B, Coupland G, Putterill J (1999) *GIGANTEA*: a circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. Embo J 18: 4679-4688

Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch S. (2005) Genetic structure and diversity in Oryza sativa L. Genetics 169(3):1631-1638

Gao L, Innan H (2008) Nonindependent domestication of the two rice subspecies, *Oryza sativa* ssp. *indica* and ssp. *japonica*, demonstrated by multilocus microsatellites. Genetics 179:965–76

Hayama R, Izawa T, Shimamoto K (2002) Isolation of rice genes possibly involved in the photoperiodic control of flowering by a fluorescent differential display method. Plant Cell Physiol 43: 494-504

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K (2003) Adaptation of photoperiodic control pathways produces short-day flowering in rice. Nature 422: 719-

722

Hayama R, Coupland G (2004) The Molecular Basis of Diversity in the Photoperiodic Flowering Responses of *Arabidopsis* and Rice. Plant Physiol 135: 677-684

Izawa T, Oikawa T, Sugiyama N, Tanisaka T, Yano M, Shimamoto K (2002) Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering in rice. Genes Dev 16: 2006-2020

Izawa T, Takahashi Y, Yano M (2003) Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and *Arabidopsis*. Curr Opin Plant Bio 6: 113-120

Izawa T (2007a) Adaptation of flowering time by natural and artificial selection in *Arabidopsis* and rice. J Exp Bot 58: 3091-3097

Izawa T (2007b) Daylength measurements by rice plants in photoperiodic short-day flowering. Int Rev Citology 256: 191-222

Kardailsky I, Shukla VK, Anh JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ, Wiegel D (1999) Activation tagging of the floral inducer FT. Science 286: 1962-1965

Kim SL, Lee S, Kim HJ, Nam HG, An G (2007) *OsMADS51* is a short-day flowering promoter that functions upstream of *EHD1*, *OsMADS14*, and *HD3A*. Plant Physiol 145: 1484-1494

Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T (1999) A pair of related genes with antagonistic roles in mediating flowering signals. Science 286: 1960-1962

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M (2002) *HD3A*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *HD1* under short-day conditions. Plant Cell Physiol 43: 1096-1105

Komiya R, Ikegami A, Tamaki S, Yokoi S, Shimamoto K (2008) *HD3A* and *RFT1* are essential for flowering in rice. Development 135: 767-774

Khush GS (1997) Origin, dispersal, cultivation and variation of rice. Plant Mol Biol. 35:25–34

Lagercrantz U (2009) At the end of the day: a common molecular mechanism for photoperiod responses in plants? J Exp Bot. 139: 1-15

Lee S, Kim J, Han JJ, Han MJ, An G (2004) Functional analyses of flowering time gene OsMADS50, the putative SUPPRESSOR OF OVEREXPRESSION OF CO 1/AGAMOUS-LIKE 20 (SOC1/AGL20) ortholog in rice. Plant Journal 38: 754-764

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and $2^{-\blacktriangle\blacktriangle C_T}$ method. Methods 25:402-408

Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of

cultivated rice, *Oryza sativa*. Proc Natl Acad Sci 103:9578–83

Ma J, Bennetzen JL (2004) Rapid recent growth and divergence of rice nuclear genomes. Proc Natl Acad Sci 101:12404–12410

Maas LF, McClung A, Moon S, Tung CW, McCouch S (2009) Dissection of a QTL reveals adaptive, interacting gene complex associated with transgressive variation for flowering time and yield in rice. Submitted to Theor. Appl. Genet.

Matsubara K, Yamanouchi U, Wang Z, Minobe Y, Izawa T, Yano M (2008) *Ehd2*, a rice ortholog of the maize *INDETERMINATE1* gene, promotes flowering by up-regulating *EHD1*. Plant Physiol 148: 1425-1435

Nanda JS (2003) Antiquity and spread of rice cultivation. In: Nanda JS and Sharma SD, editors. Monograph on Genus *Oryza*. Science Publishers, Inc. Enfield, New Hampshire, USA p. 331–346.

Onouchi H, Igeno MI, Perilleux C, Graves K, Coupland G (2000) Mutagenesis of plants overexpressing CONSTANS demonstrates novel interactions among *Arabidopsis* flowering-time genes. Plant Cell 12:885-900

Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF, Coupland G (2000) Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. Science 288: 1613-1616

Sawa M, Nusinow DA, Kay SA, Imaizumi T (2007) FKF1 and GIGANTEA complex

formation in required for day-length measurement in *Arabidopsis*. Science 318: 261-265

Sothern RB, Tseng TS, Orcutt SL, Olszewski NE, Koukkari WL (2002) *GIGANTEA* and *SPINDLY* genes linked to the clock pathway that controls circadian characteristics of transpiration in *Arabidopsis*. Chronobiol Int 19: 1005-1022

Suarez-Lopez P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G (2001) *CONSTANS* mediates between the circadian clock and the control of flowering in *Arabidopsis*. Nature 410: 1116-1120

Tadege M, Sheldon CC, Helliwell CA, Updaha N, Dennis ES, Peacock WJ (2003) Reciprocal control of flowering time by *OsSOC1* in transgenic *Arabidopsis* and by *FLC* in transgenic rice. *Plant Biotechnology Journal* 1:361-369.

Takahashi Y, Teshima KM, Yokoi S, Innan H, Shimamoto K (2009) Variations in HD1 proteins, *HD3A* promoters and *EHD1* expression levels contribute to diversity of flowering time in cultivated rice. PNAS 106: 4555-4560.

Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K (2007) HD3A Protein Is a Mobile Flowering Signal in Rice. Science 316: 1033-1036

Thomson MJ, Tai TH, McClung AM, Lai XH, Hinga ME, Lobos KB, Xu Y, Martinez CP, McCouch SR (2003) Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza*

rufipogon and the *Oryza sativa* cultivar Jefferson. Theoretical and Applied Genetics 107: 479-493

Thomson MJ, Edwards JD, Septiningsih EM, Harrington SE, McCouch SR (2006). Substitution mapping of *dth1.1*, a flowering time QTL associated with transgressive variation in rice, reveals multiple sub-QTL. Genetics 172: 2501-2514

Tseng TS, Salome PA, McClung CR, Olszewski NE (2004) *SPINDLY* and *GIGANTEA* Interact and Act in *Arabidopsis* thaliana Pathways Involved in Light Responses, Flowering, and Rhythms in Cotyledon Movements. Plant Cell 16: 1550-1563

Tsuji, H, Tamaki, S, Komiya, R, Shimamoto, K (2008) Florigen and the photoperiodic control of flowering in rice. Rice 1, 25-35

Udvardi MK, Czechowski T, Scheible WR (2008) Eleven golden rules of quantitative rt-pcr. *Plant Cell* 20:1736-1737

Vaughan DA, Balázs E, Heslop-Harrison JS (2007) From crop domestication to super domestication. Ann Bot. 100 (5):893-901

Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G (2004) Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. Science 309: 1056-1059

Vitte C, Ishii T, Lamy F, Brar D, Panaud O (2004) Genomic paleontology provides

evidence for two distinct origins of Asian rice (*Oryza sativa* L.). Mol Genet Genomics 272:504–511

Wigge PA, Kim MC, Jaegger KE, Busch W, Schmid M, Lohmann JU, Weigel D (2005) Integration of spatial and temporal information during floral induction in *Arabidopsis*. Science 309: 1056-1059

Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L, Zhou H, Yu S, Xu C, Li X, Zhang Q (2008) Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. Nature Genetics 40: 761-767

Yamamoto T, Yonemaru J, Yano M (2009) Towards the Understanding of Complex Traits in Rice: Substantially or Superficially? DNA Research 10:1-14

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) *HD1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. Plant Cell 12: 2473-2484

Yanovsky MJ, Kay SA (2003) Living by the calendar: how plants know when to flower. Nat. Rev. Mol. Cell. Biol. 4: 265- 275

CHAPTER 4

VARIATION IN FLOWERING TIME AMONG 46 DIVERSE LANDRACES OF RICE

Abstract

Forty six diverse landraces of rice representing all five major sub-populations of *O. sativa* (*indica* (12), *aus* (5), *tropical japonica* (14), *temperate japonica* (12) and *aromatic* (3), were evaluated for flowering time in growth chambers under short day (10:14h L/D) and long day (14:10h L/D) conditions. Complete sequence, including the ORF and promoter regions from two flowering time genes involved in the photoperiod pathway, *GIGANTEA* and *HEADING DATE 1,* were analyzed to look for associations between genotype and phenotype. We observed significant differences in flowering time and photoperiod sensitivity (DTF^{LD}-DTF^{SD}) among the sub-populations, were *temperate japonica* varieties flowered significantly earlier under both SD and LD conditions and presented the least sensitivity to changes in photoperiod, while *aromatic* varieties flowered later and had the highest degree of photoperiod sensitivity than any other *O. sativa* sub-population evaluated. At the same time, the range of variation was greatest within the *indica* and *aus* sub-populations. While the number of plants in the study was too small to detect significant associations between sequence variation and flowering time, we summarize our findings to provide a starting point for future association studies using the candidate gene approach.

Introduction

Flowering time is measured as the number of days from germination to anthesis. It marks the transition from the vegetative to the reproductive phase in plants. The trait is an important component of adaptation in plants and it has several pleitropic effects on other developmental traits and ultimately on yield. Genes involved in flowering time are part of a complex network of pathways that are affected by both internal factors such as plant growth factors and physiological development of plant meristems and external environmental factors, such as light, light spectrum composition and temperature.

From the characterization of genes under specific environmental conditions four distinctive and interacting pathways have been identified: (1) the photoperiod or light-dependant pathway, (2) the vernalization or temperature-dependant pathway, (3) the gibellic acid and (4) the autonomous pathway. Of these factors, light is the most predictable in a given environment, and has evolved as a major target of selection during the radiation of flowering plants (Blazquez, 2000).

The photoperiod pathway has been extensively characterized in three model systems: *Arabidopsis thaliana,* a strictly long-day plant with a vernalization requirement, *Pharbitis nil*, a strictly short-day plant, and *Oryza sativa,* a facultative short-day plant (Cremer and Coupland, 2003; Izawa, 2007; Hayama et al 2003). Characterization of natural and induced mutations and QTL studies have lead to the discovery, mapping and cloning of genes with significant roles in flowering time and provided the base for comparative studies showing that the photoperiodic pathway components (photoreceptors, circadian clock and output genes) and genes are highly conserved in plantsand that most of the differences in the genetic control of flowering are due to differential control of these essential genes (Cremer and Coupland, 2003;

Hayama and Coupland, 2004; Izawa et al. 2003; Izawa, 2007; Hayama et al 2003).

In *Arabidopsis,* the control of flowering time in response to photoperiod is mainly due to the expression of two genes, *Constans* (*CO*) and flowering time locus-T (*FT*) (Cremer and Coupland, 2003). However, *CO* expression is directly dependent on the expression of the circadian clock gene *GIGANTEA* (*GI*). The expression of these two genes is positively correlated, peaking at dawn, and together they act as promoters of flowering (Araki and Komeda, 1993). Under long-day inductive conditions, *GI* drives an increase in the expression of *CO* and the CO protein interacts with the *FLOWERING LOCUS T* (*FT)* promoter, driving its expression (Araki *et al.* 1998). *FT* is expressed in mature leaves and the protein ("florigen"*)* is transported to the shoot apical mersitem (SAM) where it interacts with *FLOWERING LOCUS T* D (*FD)* to elicit the change from vegetative growth to reproductive phase (Abe *et al.* 2005).

In rice the control of flowering time is more complex than that of *Arabidopsis* as it flowers under both short-day and long-day photoperiods. Nonetheless, the same genes are involved. *HEADING DATE 1* (*HD1*) and *HEADING DATE 3a* (*HD3A*) in rice, orthologues to *Arabidopsis CO* and *FT*, respectively, are known to play essential roles in the flowering response (Yano *et al.* 2000 and Hayama and Coupland*,* 2004). The main difference between rice and *Arabidopsis* is that *HD1* has a dual function in rice, acting as a promoter of flowering under inductive, short-day conditions and a repressor under long-day conditions (Hayama and Coupland*,* 2004). In addition, the gene *EARLY HEADING DATE 1* (*EHD1*) in rice acts as an independent inductive pathway of *HD3A* exclusively under short-day conditions (Doi *et al.* 2004). At the same time, both *HD1* and *HD3A* are members of gene families with more members in rice than in *Arabidopsis*. In rice, over 20 *COL-like* genes and at least 13 *FT-Like* genes have been identified, indicating duplication events in rice that occurred after divergence from their last common ancestor (Izawa, 2007).

Plant breeders have exploited the variation in flowering time to reduce the juvenile phase in many annual and perennial crop species (fruit and other woody species) compared to their wild relatives. The ability to manipulate flowering time helps expand the area of cultivation beyond the original center of diversity and to increase the density of plants per area of cultivation by improving competition for light while simultaneously changing plant architecture (Izawa, 2007). At the same time, both natural and artificial selection for variation in flowering time contribute to a plant's strategy for avoiding abiotic and biotic stresses, such as seasonal drought, flooding and fungal pathogens.

Genetic and phenotypic diversity provide plant breeders with the essential building blocks they need to modify a trait. In rice and other cereals, there is plenty of variation for flowering time and although the process is relatively well characterized at the molecular level,, breeders have been reluctant to make predictions based on experimental populations, due to the large pleitropic effects of flowering time on yield and grain quality, and the unpredictability of flowering models in a highly variable field environment. Early flowering in cereals is commonly associated with a reduction of yield due reduced accumulation of dry matter, but several studies in rice indicate that early flowering and high yield are not exclusive (Maas *et al.* 2009).

Association or linkage disequilibrium mapping offers a powerful approach to identify molecular polymorphisms associated with a trait of interest (Lewotin and Kojima, 1960) and has been successfully applied in numerous studies related to plant genetics (Agrama et al. 2007; Gonzalez-Martinez *et al.* 2007; Stracke et al. 2009; Sulpice et al. 2007; Wilson *et al.* 2004; Zhu et al, 2008). Association mapping offers higher mapping resolution than an equivalent number of individuals in a segregating population derived from a biparental cross in In populations with a high rate of historical recombination (Nordborg and Tavare, 2002). The statistical power of

association analysis is determined by the extent of LD, the sample size and diversity of the population and the marker density used for the study (Ersoz et al. 2007: Yu et al. 2006; Wang and Rannala, 2005). The extent of LD in different sub-populations within a species and the distribution of LD across a genome are both highly variable. This variation often results from a founder effect followed by genetic drift that leads to an unequal number of effective recombination events in any given sub-population or region of the genome (Ersoz *et al.* 2007). Furthermore, selfing plays an important role in determining the number of effective recombination events detected in a population (Nordborg, 2000). To date, there are few studies that investigate LD in rice (Chen and McCouch, 2002; Agrama, 2007; Gupta et al. 2005). The first was a targeted study of LD at a disease resistance locus (*xa5*) where it was reported that LD extends to approximately 100 kb in the *aus* supbopulation (Garris *et al.* 2003) and even further at the *waxy* domestication locus (Olsen *et al.* 2006; Agrama et al. 2007, Rakshit et al. 2007; Mather et al. 2006; McNally et al. 2009)

Sub-population structure is an important factor to consider in association studies where the relatedness of individuals in a population creates genome-wide linkage disequilibrium between unlinked loci (Gao et al. 2002; Garris et al., 2005; Lu et al. 2005; Semon et al., 2005; Ersoz *et al.* 2007). When allele frequencies between sub-populations are significantly different due to factors such as genetic drift, domestication or selection, genetic loci that do not affect whatsoever the trait may demonstrate statistical significance resulting in false positive associations (Ersoz et al. 2007). This can be accounted for using the mixed model as the basis for whole genome association mapping (Yu et al, 2006).

Where whole genome association mapping is not possible, studies can be carried out using a candidate gene approach (Andersen et al. 2005; Breseghello and Sorrells, 2006; Buckner et al. 1990; Golding, 1984; Harjes et al. 2008; Holte et al.

1997; Long et al. 1999; Spielman et al. 1994; Shaw et al. 1998; Tracy et al. 2006; Wall and Pritchard, 2003). A priori knowledge of the genes, pathways or protein processing mechanisms involved in a trait can suggest a set of target candidate genes that can be tested using an association mapping approach (Andersen et al. 2005; Breseghello and Sorrells, 2006; Buckner et al. 1990; Harjes et al. 2008; Long et al. 1999; Olsen et al. 2006; Palaisa et al. 2004; Szalma et al. 2005; Thornsberry et al. 2001). Information about allelic diversity and gene effect derived from the analysis can then be translated for use in a breeding program to improve a trait of interest.

In this study, a diverse panel of 46 rice varieties and land races were evaluated for flowering time under short-day (10 h) and long-day (14 h) photoperiods, laying the foundation of for future association studies of flowering time in rice. Strong candidate genes for the association include *HD1*, *HD3a*, *RFT1*, *OsSOC1* and *EHD1*, all of which have been shown to be sources of variation in flowering time.

Materials and Methods

Plant materials: 46 diverse lines including 5 *aus*, 3 *aromatic*, 12 *indica*, 14 *tropical japonica* and 12 *temperate japonica* were evaluated for flowering time in 2007. Plant materials used in this experiment are summarized on Table 4.1 and are part of the diversity panel utilized by Garris et al. (2003) and Caicedo et al., (2007).

Growth chamber experiments: Seeds were sown in pots in August/2007 and March/2008 in the Guterman greenhouse at Cornell University and 21 day old seedlings were transplanted to 50-mm-wide x 178-mm deep plastic pots and grown under neutral day-length (12 h of sunlight) at 33/28 C (day/night). At 24 days of age, plants were transferred to growth chambers (Conviron Pembina, ND) where lighting

and environmental conditions could be maintained constant throughout the growing season. In the growth chambers, half of the plants received a short day treatment (10:14 L/D) and the other half received a long day treatment (14:10 L/D). Light intensity was 10,000 lx, humidity was 70%, daytime temperature was 28 ^{o}C d and the night temperature was 25 ^{o}C. A total of 14 plants per line, organized in a randomized complete block design (RCBD), were grown in each growth chamber and all plants were under-watered at a constant water level.

Phenotypic evaluation: The following traits were evaluated in both controlled environments (1) days to flowering (DTF) defined as the number of days from seedling emergence until 50% of the main tillers had spikelets with extruded stigmas and (2) Photoperiod sensitivity defined by the difference of flowering time under long-day minus flowering time under short-day as a measure of flowering stability among the lines evaluated.

Statistical Analysis: Analysis of variance for flowering time was performed using the JMP statistical package, version 7.0 for Windows (SAS Institute Inc., Cary, NC). Data from all experiments were normalized by eliminating extreme values and all assumptions of the Least-Square model (LSM) were tested for growth chamber data to estimate the variance components. The LSM included the following fixed effects: genotypes (G), Growth Chambers (E), genotypes nested within photoperiod (G(PP)) and genotype by environment interaction (G x E). Multiple means comparisons of all lines for flowering time and yield were done using the Student T ($p<0.0001$).

Results

Flowering time under growth chamber conditions: Flowering time was measured as the number of days to flowering after germination of seedlings. The earliest flowering line was Geumobyeo which flowered at 51 and 66 days under SD and LD photoperiods, respectively (Table 4.1 and Figure 4.1). The latest flowering lines included BJ-1 and Basmati-217 that flowered at 114 and 120 days under SD and 132 and 137 days under LD, respectively (Table 4.1 and Figure 4.1).

Table 4.1 Days to flowering (DTF) of 46 diverse rice lines under short-day and long-day photoperiods. Photoperiod sensitivity is defined as the difference between DTF^{LD}-DTF^{SD}.

		Long Day		**Short Day**		**PS***
Line	**Sub-Population**	**DTF**	**Std Dev**	**DTF**	**Std Dev**	**LD-SD**
93-11	*Indica*	124.2	17.4	96.8	7.6	27.4
Ai Chiao hong	*Indica*	72.5	7.5	75.5	9.9	-3.0
Azucena	*Tropical Jap.*	107.1	11.1	91.2	4.0	15.9
Basmati-217	*Aromatic*	137.3	13.7	120.3	10.0	17.0
Binulawan	*Indica*	78.0	7.4	70.3	9.8	7.7
BJ-1	*Aus*	131.8	20.8	114.0	2.3	17.8
Chau	*Indica*	128.6	4.8	82.2	16.9	46.4
Davao	*Tropical Jap.*	129.7	10.4	53.8	3.1	75.9
Dee geo woo gen	*Indica*	131.6	20.4	107.0	6.6	24.6
Dhala Shaitta	*Aus*	81.7	16.6	58.1	3.3	23.6
Dom-Sofid	*Aromatic*	99.1	18.4	83.5	6.2	15.5

Table 4.1 (continue)

DV85	*Aus*	100.9	21.1	88.4	8.6	12.5
Early Wateribune	*Temperate Jap.*	75.7	8.5	52.8	5.3	22.9
Fortuna	*Tropical Jap.*	132.5	5.7	98.7	35.3	33.8
Geumobyeo	*Temperate Jap.*	66.1	6.6	50.8	3.9	15.2
Gotak-Gatik	*Tropical Jap.*	71.8	6.5	68.0	10.5	3.8
Guan Yin Tsan	*Indica*	90.2	14.1	98.3	9.6	-8.2
Gyehura-3	*Temperate Jap.*	94.0	31.6	67.3	4.5	26.7
Hsia Chioh Keh Tu	*Indica*	130.9	13.3	88.2	12.2	42.7
IR64	*Indica*	97.0	20.1	75.7	12.5	21.3
Jambu	*Tropical Jap.*	98.3	20.9	81.5	9.1	16.8
Jefferson	*Tropical Jap.*	94.0	36.1	66.2	6.7	27.8
Jhona-349	*Aus*	105.4	13.6	97.1	7.5	8.3
Kalukantha	*Indica*	137.0	0.0	86.5	16.3	50.5
Khao Hawn	*Tropical Jap.*	100.5	13.6	97.0	8.7	3.5
Koshikari	*Temperate Jap.*	71.0	2.5	67.0	11.4	4.0
Kotobuki Mochi	*Tropical Jap.*	88.2	14.4	73.9	8.9	14.3
KU115	*Tropical Jap.*	114.0	19.7	90.0	3.7	24.0
Lemont	*Tropical Jap.*	86.9	19.2	65.6	5.2	21.2
Luk Takhat	*Temperate Jap.*	99.0	17.2	71.9	8.0	27.1
M202	*Temperate Jap.*	80.2	12.8	69.6	3.4	10.6
Mansaku	*Temperate Jap.*	81.1	18.9	70.2	7.2	10.9
Mirti	*Tropical Jap.*	69.3	8.3	56.9	5.4	12.4
Mudgo	*Indica*	107.2	14.9	74.3	11.3	33.0
Nipponbare	*Temperate Jap.*	86.8	10.5	51.2	4.9	35.6
NPE-844	*Tropical Jap.*	105.3	22.2	87.7	18.1	17.6

Table 4.1 (continue)

Padi Kasalle	*Tropical Jap.*	137.0	0.0	105.0	0.0	32.0
Pankhari-203	*Aromatic*	140.7	6.4	90.1	13.7	50.5
Phudugey	*Aus*	104.8	24.1	77.3	7.6	27.5
Rathuwee	*Indica*	129.4	11.0	93.0	12.7	36.4
Shinriki	*Temperate Jap.*	89.2	14.8	69.0	14.0	20.2
Ta Hung ku	*Temperate Jap.*	69.8	3.2	55.6	2.6	14.3
Taducan	*Indica*	101.8	2.6	86.4	13.5	15.3
Tequing	*Indica*	104.9	17.1	95.2	5.7	9.7
Trembese	*Tropical Jap.*	97.4	17.2	88.1	9.0	9.3
WC-6	*Temperate Jap.*	78.7	2.6	79.0	7.3	-0.3

The lines were grouped into four categories under SD conditions based on multiple mean comparisons: early flowering (<60 days), intermediate or normal flowering (60-75 days), late flowering (75-90) and extremely late flowering (>90) (Table 4.1). Early lines included: Geumobyeo, Nipponbare, Early Wateribune, Davao, Ta Hung Ku, Mirti and Dhala Shaitta. Extremely late lines included: Basmati-217, BJ-1, Dee Geo Woo Gen and Padi Kasalle all of which flowered later than 100 days (Dunnet's multiple mean comparison ($P<0.001$); Table 4.1).

Statistical analysis indicates that there is significant variation in flowering time among the five *O. sativa* sub-populations. Overall the average flowering time per sub-population under SD and LD is as follows: *temperate japonica* = av 64 (SD) and 81 (LD) days <*tropical japonica* = av 77 (SD) and 101 (LD) days < *aus* = av 80 (SD) and 106 (LD) days < *indica* = av 85 (SD) and 109 (LD) days and <*aromatic* = av 86 (SD) and 118 (LD) days, respectively (Table 4.2 and Figure 4.2). Under both SD and LD conditions *temperate japonicas* were significantly earlier than the rest of sub-

Figure 4.1 Days to flowering comparison of 46 diverse rice lines under growth chamber conditions. Lines are color coded to distinguish distribution of flowering within each sub-population: *temperate japonicas* (blue), *tropical japonicas* (Red), *indica* (Orange), *aus* (Light Green) and *aromatics* (Yellow).

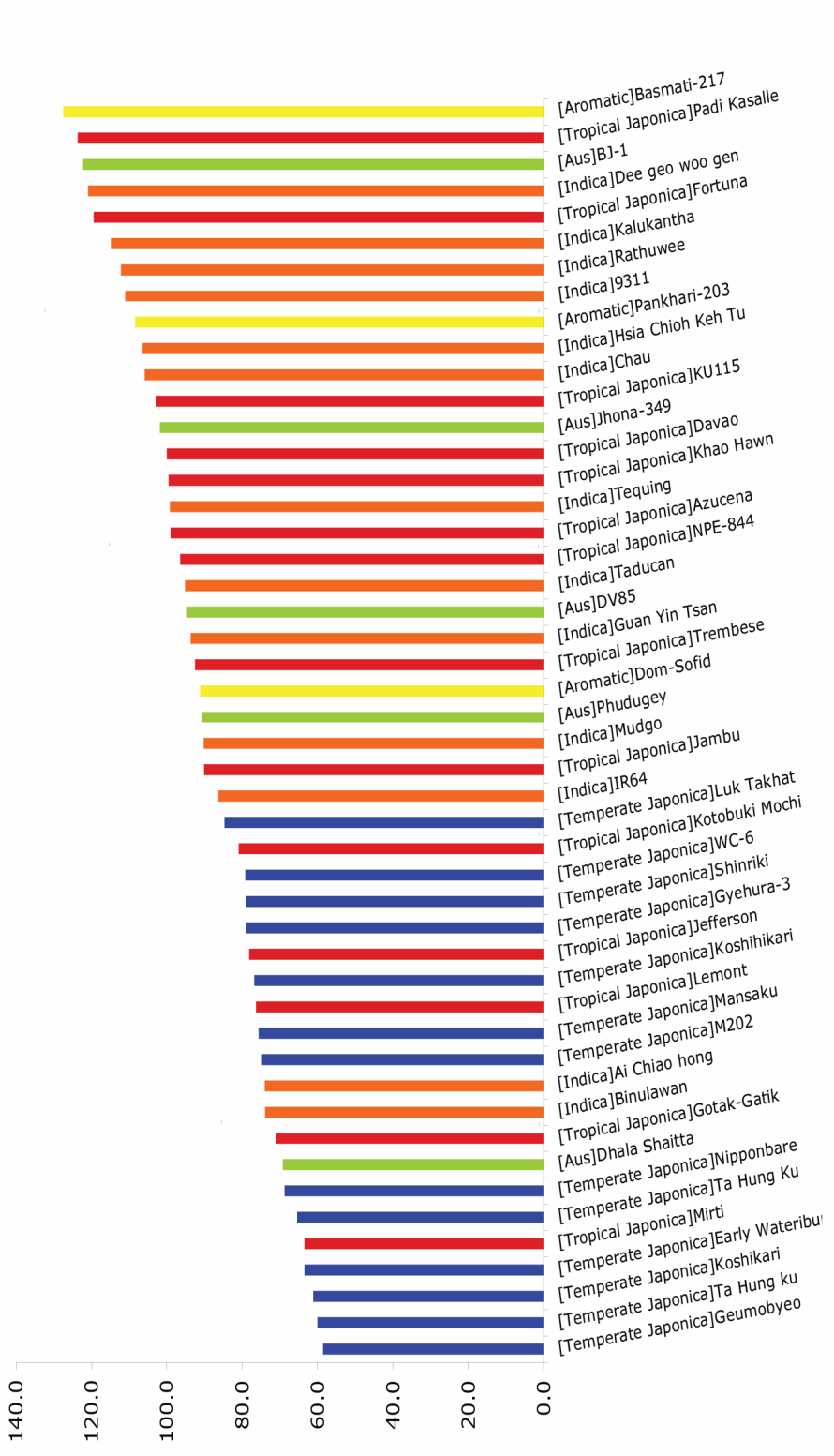
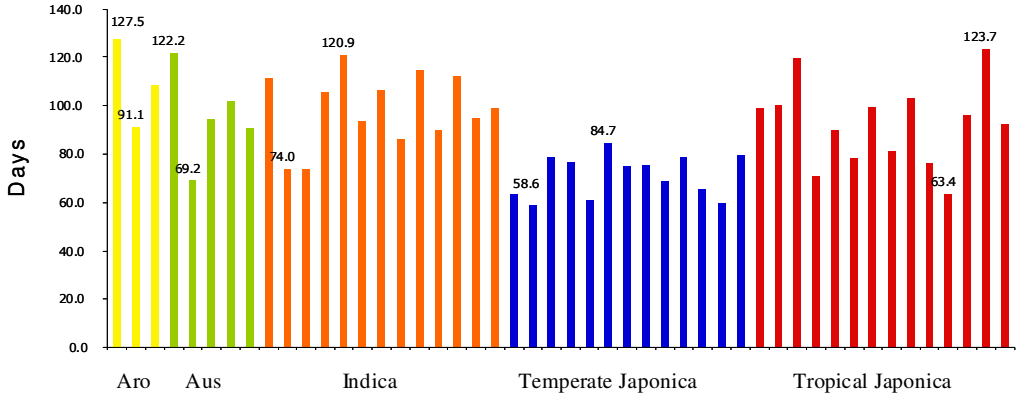
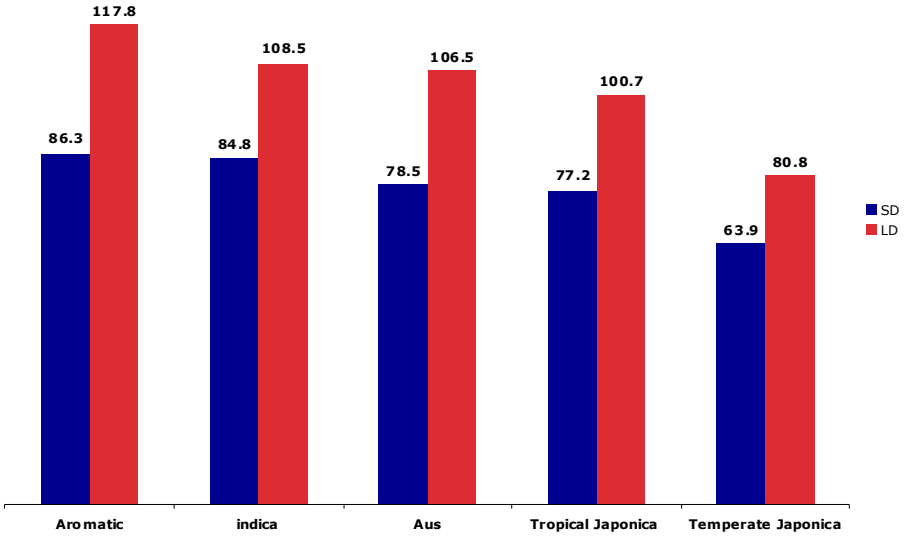

[Aromatic]Basmati-217
[Tropical Japonica]Padi Kasalle
[Aus]BJ-1
[Indica]Dee geo woo gen
[Tropical Japonica]Fortuna
[Indica]Kalukantha
[Indica]Rathuwee
[Indica]9311
[Aromatic]Pankhari-203
[Indica]Hsia Chioh Keh Tu
[Indica]Chau
[Tropical Japonica]KU115
[Aus]Jhona-349
[Tropical Japonica]Davao
[Tropical Japonica]Khao Hawn
[Indica]Tequing
[Tropical Japonica]Azucena
[Tropical Japonica]NPE-844
[Indica]Taducan
[Aus]DV85
[Indica]Guan Yin Tsan
[Tropical Japonica]Trembese
[Aromatic]Dom-Sofid
[Aus]Phudugey
[Indica]Mudgo
[Tropical Japonica]Jambu
[Indica]IR64
[Temperate Japonica]Luk Takhat
[Tropical Japonica]Kotobuki Mochi
[Temperate Japonica]WC-6
[Temperate Japonica]Shinriki
[Temperate Japonica]Gyehura-3
[Tropical Japonica]Jefferson
[Temperate Japonica]Koshihikari
[Tropical Japonica]Lemont
[Temperate Japonica]Mansaku
[Temperate Japonica]M202
[Indica]Ai Chiao hong
[Indica]Binulawan
[Tropical Japonica]Gotak-Gatik
[Aus]Dhala Shaitta
[Temperate Japonica]Nipponbare
[Temperate Japonica]Ta Hung Ku
[Tropical Japonica]Mirti
[Temperate Japonica]Early Wateribu
[Temperate Japonica]Koshikari
[Temperate Japonica]Ta Hung ku
[Temperate Japonica]Geumobyeo
140.0
120.0
100.0
80.0
60.0
40.0
20.0
0.0

Figure 4.2 (A) Days flowering distribution within each sub-population. Lines included 3 *aromatic* lines, 5 *aus*, 12 *indicas*, 12 *temperate japonicas* and 14 *tropical japonicas*. (B) Sub-population comparison under SD (dark blue bars) and LD (red bars) conditions indicates that Temperate Japonicas are earlier flowering while Aromatics are the latest flowering group.

A


Days
140.0
120.0
100.0
80.0
60.0
40.0
20.0
0.0
127.5
91.1
122.2
69.2
74.0
120.9
58.6
84.7
63.4
123.7
Aro
Aus
Indica
Temperate Japonica
Tropical Japonica

B


86.3
117.8
84.8
108.5
78.5
106.5
77.2
100.7
63.9
80.8
SD
LD
Aromatic
indica
Aus
Tropical Japonica
Temperate Japonica

populations evaluated and had the lowest photoperiod sensitivity (PS) (Dunnet's Test P<0.0001). In contrast, lines of the *aromatic* sub-population flower the latest and were the most sensitive to changes in photoperiod. *Tropical japonicas*, *indicas* and *aus* similar to each other and intermediate to *temperate japonicas* and *aromatics*.
The range of flowering time within the *temperate japonica* sub-population was significantly less than for any of the other tropical-adapted sub-populations, and it was greatest for the *aromatic* sub-population. In all sub-populations, the range in days to flowering was consistently greater under LD (66 to 130 days) than under SD (51 to 89 days) (Figure 4.2a and b). The range of flowering for *temperate japonica* was between 51 and 84 days; for *tropical japonica* between 63 and 126; *indica* between 74 and 121; *aus* between 59 and 122 and *aromatic* between 91 to 128 (Supplementary Table 4.1 and 4.2).

Table 4.2 Mean comparison (P<0.001) of days to flowering (DTF) in *O.sativa* sub-populations under short (SD) and long day conditions (LD). Photoperiod sensitivity (PS) for each sub-population was obtained by subtracting DTF^{LD}- DTF^{SD}.

Sub-population	N	Mean LD[1]		Mean SD[1]		Mean PS[2]	
aromatic	25	117.84	C	86.32	C	31.52	C
aus	43	106.51	B	78.54	B	27.97	B
indica	145	108.49	B	84.81	C	23.68	B
tropical Japonica	144	100.72	B	77.23	B	23.49	B
temperate Japonica	97	80.75	A	63.89	A	16.86	A

[1] Dunnet's multiple mean comparison (P<0.001) sub-populations with different letter significantly earlier.
[2] Photoperiod sensitivity (PS)= Mean LD – Mean SD

Under short-days, these lines flower an average of 24 days earlier than under long-day conditions, indicating a high degree of photoperiod sensitivity in most lines.

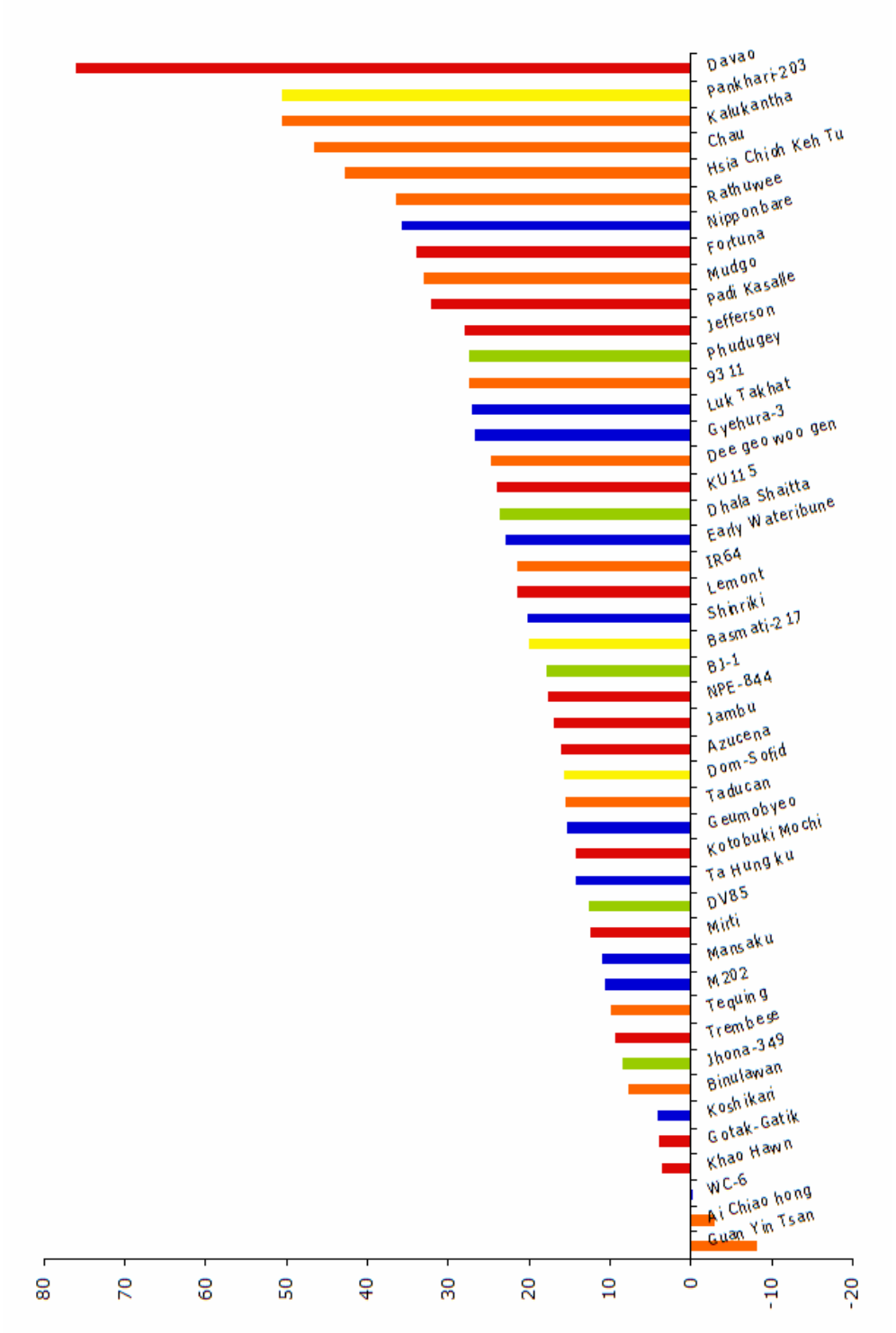
The level of photoperiod sensitivity was significantly less ($p<0.001$) among *temperate japonica* varieties<compared to *indica*, *aus, tropical japonica* < and aromatic varieties (Figure 4.3). Interestingly, two *indica* lines, Ai Chiao Hong and Guan Yin Tsan, showed the reverse effect whereby they flowered 8 and 6 days earlier under LD than SD. These varieties are worthy of further study.

Discussion

Flowering time is one of the most intensively studied physiological phenomena in plants and the genetic control of the trait has intrigued scientist for many years. Numerous genes have been cloned and flowering time mutants have been characterized to elucidate the underlying molecular mechanisms involved in the regulation of flowering time.

Plant breeders have benefited from these studies by identifying gene targets and sources of variation that could be used in marker assisted selection to develop lines with modified flowering time. This has enabled them to expand the area of cultivation of a high yielding variety, Koshihikari, in Japan (Takeuchi *et al.* 2006) and to enhance earliness and yield in a US *tropical japonica* cultivar, Jefferson, by marker-assisted backcrossing to a wild relative, *O. rufipogon* (Maas *et al.* 2008 and Thompson *et al.* 2006). The limitation of trying to manipulate flowering time by targeting specific genes in breeding programs is due to the extensive pleitropic effects that accompany changes in flowering time and tend to affect yield and yield components in different environments. The use of association mapping promises to help expand the repertoire of genes, alleles and genetic networks that can be targeted by plant breeders and to enhance the efficiency of marker-assisted and genomic selection strategies in the development of new crop varieties.

Figure 4.3 Photoperiod sensitivity comparison of 46 diverse rice lines under growth chamber conditions. Lines are color coded to distinguish distribution of flowering within each sub-population: *temperate japonicas* (blue), *tropical japonicas* (Red), *indica* (Orange), *aus* (Light Green) and *aromatics* (Yellow).


Davao
Pankhari-203
Kalukantha
Chau
Hsia Chich Keh Tu
Rathuwee
Nipponbare
Fortuna
Mudgo
Padi Kasalle
Jefferson
Phudugey
9311
Luk Takhat
Gyehura-3
Dee geo woo gen
KU115
Dhala Shaitta
Early Watericune
IR64
Lemont
Shinriki
Basmati-217
BJ-1
NPE-844
Jambu
Azucena
Dom-Sofid
Taducan
Geumobyeo
Kotobuki Mochi
Ta Hungku
DV85
Miri
Mansaku
M202
Tequing
Trembese
Jhona-349
Binulawan
Koshikari
Gotak-Gatik
Khao Hawn
WC-6
Ai Chiao hong
Guan Yin Tsan
80
70
60
50
40
30
20
10
0
-10
-20

Our study is limited by the small number of lines evaluated for flowering time. The number of lines required for an association depends largely on the extent of LD and population sub-structure. In maize a total of 1000 genotypes is the minimum requirement for significant association between haplotypes and the trait of interest to be detected (Personal communication William Beavis). In the future, this data may be augmented by evaluating a larger sample of rice lines from the diversity panel and a larger set of candidate flowering time genes, such as *HD1*, *HD3a*, *RFT1*, *OsSOC1* and *EHD1,* all of which have been shown to have a significant impact on flowering time.

Acknowledgements

We would like to thank Fumio Onishi and Jean Koski in the Plant Breeding and Genetics Department at Cornell University for their assistance in the management of the rice populations used in this study. Funding for this research was provided by an NSF_Plant Genome Research Program (#0606461) grant. The association analysis was a collaboration with Badri Padhukasahasram in the Bustamente research group in the Department of Biological Statistics and Computational Biology at Cornell University. We are grateful to to Greg Wilson for helping to align the sequence data for the candidate genes *GIGANTEA* and *HEADING DATE 1* in the 46 lines evaluated.

APPENDICES

Appendix Table 4.1 Mean comparison of 46 diverse lines under short-day.

Line	Statistical Significance (P<0.001)											DTF
BJ-1	A											114.0
Dee geo woo gen	A											107.0
Padi Kasalle	A	B										105.0
Fortuna	A	B	C	D								98.7
Guan Yin Tsan	A	B	C									98.3
Jhona-349	A	B	C									97.1
Khao Hawn	A	B	C									97.0
9311	A	B	C									96.8
Tequing	A	B	C	D								95.2
Rathuwee		B	C	D	E							93.0
Azucena		B	C	D	E	F						91.2
Pankhari-203			C	D	E	F	G					90.1
KU115		B	C	D	E	F	G	H				90.0
DV85				D	E	F	G	H				88.4
Hsia Chioh Keh Tu				D	E	F	G	H				88.2
Trembese				D	E	F	G	H				88.1
NPE-844				D	E	F	G	H				87.7
Kalukantha			C	D	E	F	G	H	I	J	K	86.5
Taducan					E	F	G	H				86.4
Dom-Sofid						F	G	H	I			83.5
Chau							G	H	I	J		82.2
Jambu								H	I	J		81.5

Appendix Table 4.1 (continue)

WC-6	I	J	K	L						79.0
Phudugey	I	J	K	L	M					77.3
IR64		J	K	L	M					75.7
Ai Chiao hong		J	K	L	M					75.5
Mudgo			K	L	M	N				74.3
Kotobuki Mochi			K	L	M	N				73.9
Luk Takhat				L	M	N	O			71.9
Binulawan					M	N	O			70.3
Mansaku					M	N	O			70.2
M202					M	N	O			69.6
Shinriki					M	N	O			69.0
Gotak-Gatik						N	O			68.0
Gyehura-3						N	O			67.3
Koshihikari						N	O			67.0
Jefferson							O			66.2
Lemont							O			65.6
Dhala Shaitta								P		58.1
Mirti								P	Q	56.9
Ta Hung Ku								P	Q	55.6
Davao								P	Q	53.8
Early Wateribune								P	Q	52.8
Nipponbare								P	Q	51.2
Geumobyeo									Q	50.8

Appendix Table 4.2 Mean comparison of 46 diverse lines under long-day.

Line	Statistical Significance (P<0.001)								DTF
Pankhari-203	A	B							140.7
Basmati-217	A								137.3
Kalukantha	A	B							137.0
Padi Kasalle	A	B							137.0
Fortuna	A	B							132.5
BJ-1	A	B							131.8
Dee geo woo gen	A	B							131.6
Hsia Chioh Keh Tu	A	B							130.9
Davao	A	B							129.7
Rathuwee	A	B							129.4
Chau	A	B							128.6
9311		B	C						124.2
KU115			C	D					114.0
Mudgo				D	E				107.2
Azucena				D	E				107.1
Jhona-349				D	E				105.4
NPE-844				D	E				105.3
Tequing				D	E				104.9
Phudugey				D	E				104.8
Taducan				D	E	F	G	H	101.8
DV85					E	F			100.9
Khao Hawn					E	F			100.5
Dom-Sofid					E	F	G		99.1
Luk Takhat					E	F	G	H	99.0

Appendix Table 4.2 (continue)

Jambu	E	F	G	H					98.3
Trembese	E	F	G	H					97.4
IR64	E	F	G	H					97.0
Gyehura-3	E	F	G	H	I				94.0
Jefferson	E	F	G	H	I	J			94.0
Guan Yin Tsan		F	G	H	I				90.2
Shinriki		F	G	H	I	J			89.2
Kotobuki Mochi			G	H	I	J			88.2
Lemont				H	I	J			86.9
Nipponbare				H	I	J			86.8
Dhala Shaitta					I	J	K		81.7
Mansaku					I	J	K		81.1
M202					I	J	K		80.2
WC-6					I	J	K		78.7
Binulawan						J	K	L	78.0
Early Wateribune					I	J	K	L	75.7
Ai Chiao hong							K	L	72.5
Gotak-Gatik							K	L	71.8
Koshikari		F	G	H	I	J	K	L	71.0
Ta Hung ku							K	L	69.8
Mirti							K	L	69.3
Geumobyeo								L	66.1

Appendix Table 4.3 ANOVA of Days to flowering (DTF) under growth chambers.

Variance Component	**DTF in Control Environments**			
	SS	% Total	F Value	Prob (F)
δ^2_G	41199.3	26.78	123.3	<.0001
δ^2_{PP}	74699.5	48.55	4249.0	0.0001
$\delta^2_{R(PP)}$	4705.3	3.06	7.9	<.0001
δ^2_{GXPP}	6320.4	4.11	18.9	<.0001

REFERENCES

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, and Araki T (2005) *FD*, a *bZIP* Protein Mediating Signals from the Floral Pathway Integrator FT at the Shoot Apex. Science 309: 1052

Agrama, H. A. (2007) Molecular diversity and genome-wide linkage disequilibrium patterns in a worldwide collection of Oryza sativa and its wild relatives. *Euphytica*

Agrama, H. A., Eizenga, G. C. and Yan, W. (2007) Association mapping of yield and its components in rice cultivars. Mol. Breeding 19:341-356

Andersen JR, Schrag T, Melchinger AE, Zein I, Lubberstedt T (2005) Validation of Dwarf8 poly-morphisms associated with flowering time in elite European inbred lines of maize (Zea maysL.). Theor Appl Genet 111:206–217

Araki T, Komeda Y (1993) Analysis of the Role of the Late-Flowering Locus, Gl, in the Flowering of *Arabidopsis thaliana*. Plant Journal 3: 231-239

Araki T, Kobayashi Y, Kaya H, Iwabuchi M (1998) The flowering time gene FT and regulation of flowering in *Arabidopsis*. Journal of Plant Research 111: 277-281

Blazquez MA (2000) Flower development pathways. J Cell Sci 113:3547-3548

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165–1177

Buckner B, Kelson TL, Robertson DS (1990) Cloning of the y1 locus of maize, a gene

involved in the biosynthesis of carotenoids. Plant Cell 2:867–876

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, et al. 2007 Genome-Wide Patterns of Nucleotide Polymorphism in Domesticated Rice. PLoS Genetics 3(9):

Chen X, Cho YG, McCouch SR (2002) Sequence divergence of rice microsatellites in *Oryza* and other plant species. Mol Genet Genomics 268:331–343

Cremer F, Coupland G (2003) Distinct photoperiodic responses are conferred by the same genetic pathways in *Arabidopsis* and in rice. Trends Plant Sci 8:405-407

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M and Yoshimura A (2004) *EHD1*, a B-type response regulator in rice, confers short day promotion of flowering and controls *FT-like* gene expression independently of *HD1*. Genes & Dev 18:926-936

Eizenga, G. C. (2008) Exploring genetic diversity and potential novel disease resistance genes in a collection of rice (Oryza spp.) wild relatives. *Genetic Resources and Crop Evolution*

Ersoz ES, Yu J, Buckler ES 2007. Applications of linkage disequilibrium and association mapping in crop plants. Genomics Assisted Crop Improvement Vol. 1: Genomics Approaches and Platforms (eds. R.K. Varshney and R. Tuberosa), pp. 97-119.

Gao LZ, Schaal BA, Zhang CH, Jia JZ (2002) Assessment of population genetic structure in common wild rice *Oryza rufipogon* Griff. Using microsatellite and allozyme markers. Theor Appl Genet 106:173–180

Garris A, McCouch SR, Kresovich S (2003) Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the *xa5* locus of rice (*Oryza sativa* L.). Genetics 165:759–769

Garris A, Tai T, Coburn J, Kresovich S, McCouch S (2005) Genetic structure and diversity in *Oryza sativa* L. Genetics 169:1631–1638

Golding GB (1984) The sampling distribution of linkage disequilibrium. Genetics 108:257–274

Gonzalez-Martinez SC, Wheeler N, Ersoz ES, Neale DB (2006) Association genetics in Pintaeda L.I. Wood property traits. Genetics 175:399–409

Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Mol Biol 57:461–485

Hayama R, Izawa T, Shimamoto K (2002) Isolation of rice genes possibly involved in the photoperiodic control of flowering by a fluorescent differential display method. Plant Cell Physiol 43**:** 494-504

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K (2003) Adaptation of photoperiodic control pathways produces short day flowering in rice. Nature 422: 719-

722

Harjes CE, Rocheford TR, Bai L,, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan J, Buckler ES (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. Science 319(5861):330–333

Holte S, Quiaoit F, Hsu L, Davidov O, Zhao LP (1997) A population based family study of a common oligogenic disease – part I: association/aggregation analysis. Genet Epidemiol 14:803–807

Hayama R, Coupland G (2004) The Molecular Basis of Diversity in the Photoperiodic Flowering Responses of *Arabidopsis* and Rice. Plant Physiol 135: 677-684

Izawa T, Takahashi Y, Yano M (2003) Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and *Arabidopsis*. Curr Opin Plant Bio 6: 113-120

Izawa T (2007) Adaptation of flowering time by natural and artificial selection in *Arabidopsis* and rice. Journal Exp Bot 58: 3091-3097

Lewontin RC, Kojima K (1960) The evolutionary dynamics of complex polymorphisms. Evolution 14:458–472

Long AD, Langley CH (1999) The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. Genome Res 9:720–731

Lu H, Redus MA, Coburn JR, Rutger JN, McCouch SR, Tai TH (2005) Population structure and breeding patterns of 145 US rice cultivars based on SSR marker analysis. Crop Sci 45:66–76

Maas LF, McClung A, Moon S, Tung CW, McCouch S (2009) Dissection of a QTL reveals adaptive, interacting gene complex associated with transgressive variation for flowering time and yield in rice. Submitted to Theoretical and Applied Genetics.

Mather KA, Caicedo AL, Polato NR, Olsen KM, McCouch S, Purugganan MD. (2007) The extent of linkage disequilibrium in rice (Oryza sativa L.). Genetics 177:2223-2232.

McNally KL et al (2009) Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. PNAS 106:12273-12278

Nordborg M (2000) Linkage disequilibrium, gene trees, selfing: an ancestral recombination graph with partial self-fertilization. Genetics 154:923–929

Nordborg M, Tavare S (2002) Linkage disequilibrium: what history has to tell us. Trends Genet 18:83–90

Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan D (2006). Selection under domestication: evidence for a sweep in the rice *Waxy* genomic region. Genetics 173:975–983

Palaisa K, Morgante M, Tingey S, Rafalski A (2004). Long-range patterns of diversity and linkage disequilibrium surrounding the maize Y1 gene are indicative of an asymmetric selective sweep. PNAS 101:9885–9890

Rakshit S, Rakshit A, Matsumura H, Takahashi Y, Hasegawa Y,Ito A et al. (2007). Large-scale DNA polymorphism study of Oryza sativa and O. rufipogon reveals the origin and divergence of Asian rice. Theor. Appl. Genet. 114:731–743

Semon M, Nielsen R, Jones MP, McCouch SR. (2005). The population structure of African cultivated rice oryza glaberrima (Steud.): evidence for elevated levels of linkage disequilibrium caused by admixture with O. sativa and ecological adaptation. Genetics. 169:1639-47.

Shaw SH, Carrasquillo MM, Kashuk C, Puffenberger EG, Chakravarti A (1998) Allele frequency distributions in pooled DNA samples: applications to mapping complex disease genes. GenomeRes 8:111–123

Spielman RS,McGinnis RE, EwensWJ (1994) The transmission/disequilibrium test detects cosegregation and linkage. Am J Hum Genet 54:559–560; author reply 560–553

Stracke S, Haseneyer G, Veyrieras JB, et al. (2009) Association mapping reveals gene action and interactions in the determination of flowering time in barley. Ther. Appl. Genet. 118:259–273

Sulpice R, Gibon Y, Tschoep H, et al. (2007) Association mapping within Arabidopsis

accessions to identify regulation networks in the primary metabolism. Development of a rapid and sensitive non-radioactive assay for ribulose-1,5-bisphosphate carboxylase. Photosynthesis Res. 91: 238

Szalma SJ, Buckler ES, SnookME,McMullenMD (2005) Association analysis of candidate genesfor maysin and chlorogenic acid accumulation in maize silks. Theor Appl Genet 110:1324–1333

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler IV ES (2001). *Dwarf8* polymorphisms associate with variation in flowering time. Nat Genet 28:286–289

Tracy WF, Whitt SR, Buckler ES (2006). Recurrent mutation and genome evolution: example ofSugary1 and the origin of sweet maize. Crop Sci 46:1–7

Wall JD, Pritchard JK (2003) Haplotype blocks and linkage disequilibrium in the human genome. Nat Rev Genet 4:587–597

Wang Y, Rannala B (2005) In silico analysis of disease-association mapping strategies using the coalescent process and incorporating ascertainment and selection. Am J Hum Genet 76:10661073

Wilson LM, Whitt SR, Ibanez AM, Rocheford TR, Goodman MM, Buckler ES (2004) Dissection of maize kernel composition and starch production by candidate gene association. Plant Cell 16:2719–2733

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) *HD1*, a major photoperiod

sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. Plant Cell 12**:** 2473-2484

Yu J, Buckler ES (2006) Genetic association mapping and genome organization of maize. Curr Opin Biotech 17:155–160

CHAPTER 5

CONCLUSIONS AND FUTURE PERSPECTIVES

Conclusions

This dissertation provides the basis for understanding the genetic basis of transgressive variation associated with the flowering time QTL, *dth1.1,* in rice. We demonstrate that single, well-defined regions of introgression from the late-flowering wild relative, *O. rufipogon,* can confer earliness in the U.S. *tropical japonica* variety, Jefferson, and that differential expression of candidate flowering time genes located within the *dth1.1* QTL is predictive of flowering time in the introgression lines. We also evaluated flowering time under short day conditions in 46 diverse rice varieties as a starting point for future association mapping studies. Our aim was to address three questions: (1) Which regions of introgression within the *dth1.1* QTL from *O. rufipogon* are necessary to reduce the number of days to flowering in the Jefferson genetic background? (2) How is gene expression at key flowering time loci affected by the presence of *O. rufipogon* introgressions conferring early or late flowering in the Jefferson sub-introgression lines (SILs)? (3) What is the range of variation in flowering time observed among a set of 46 diverse *O. sativa* lines grown under short day conditions in the growth chamber?

The *dth1.1* QTL was dissected into a series of SILs that each contained a single candidate flowering time gene from *O. rufipogon.* The candidate gene provided a sentinel for each SIL, and though the lines each contained a hundred or more genes from *O. rufipogon* within each introgression, we observed that SILs conferring transgressive early flowering always contained at least two candidate flowering time sentinel genes. Lines that flowered earlier than Jefferson under SD conditions included

SIL_GI/SOC1, SIL_SOC1/FT-L8, SIL_EMF1/PNZIP and SIL_GI/SOC1/FT-L8. Lines carrying introgressions with single candidate flowering time genes flowered similarly to the recurrent parent, Jefferson.

Expression analysis of the four flowering time genes *OsGI*, *HD1, HD3A* and *RFT1* in the SILs and the two parents revealed that differences in flowering time can not only be predicted by changes in levels of expressions of these genes, as has been previously observed in mutational analysis of these genes in rice and *Arabidopsis,* but also to the regulation of gene expression over time. The maximum expression of these 4 genes in the early flowering parental line, Jefferson, occurred about 20 days earlier than in the late flowering parental line, *O. rufipogon,* under SD conditions. Gene expression in Jefferson peaked at approximately 50 days after germination (DAG) while it peaked around 90 DAG in *O. rufipogon*; in both cases, peak expression levels coincided with the transition from the vegetative to the reproductive phase. The late flowering habit of *O. rufipogon* can be attributed to an up-regulation of the flowering-time repressor gene, *OsGI,* and to down-regulation of *HD3A (OsFT).* We hypothesize that the down-regulation of *HD3A* under inductive short day conditions in *O. rufipogon* may be caused by a defective (non-functional) promoter sequence in the *O. rufipogon* allele at *HD3A*. This is supported by the late flowering phenotype of P9-84 which carries an *O. rufipogon* allele at *HD3A* on chromosome 6, reversing the effect of the early flowering caused by the *O. rufipogon* alleles at *dth1.1*.

Despite the low level of expression of *HD3A* in *O. rufipogon,* the accession eventually flowers. From this we infer that one of the other FT-family members, such as *RFT1* or one of the other 13 *FT-Like* genes that have not been extensively studied in rice, is likely to act redundantly to *HD3A* allele in *O. rufipogn* (though with a weaker effect) and is capable of eventually inducing flowering in *O. rufipogon*. Our hypothesis is supported by the late flowering phenotype of pre-SIL P9-84, which

carries *O. rufipogon* alleles in the *dth1.1* introgression on chromosome 1 that normally induce early flowering, but in addition, carries *O. rufipogon* alleles at *HD3A* and *RFT1* on chromosome 6. The early transgressive flowering phenotype observed in SIL_GI/SOC1, SIL_SOC1/FT-L8, SIL_EMF1/PNZIP and SIL_GI/SOC1/FT-L8 can be attributed mainly to an up-regulation of *HD3A,* beyond that of the recurrent parent Jefferson. The increased expression levels of *HD3A* can only be attributed to the presence of *O. rufipogon* alleles in the introgressions present in each SIL.

Results from the evaluation of flowering time variation and photoperiod sensitivity among the 46 diverse accessions of *O. sativa* showed significant differences among sub-populations: *temperate japonica* varieties tended to flower earlier and have less photoperiod sensitivity than other subpopulations, while a*romatic* varieties flowered later and had the highest degree of photoperiod sensitivity of the five groups. However, there was a wide range of variation in flowering time within each sub-population, suggesting high levels of allelic variation in the network of flowering time genes within each sub-group. Taken together these results indicate that the distinctive pattern of flowering time variation observed in the *temperate japonica* subpopulation is correlated with the extremely short length of the growing season coupled with the long daylengths and low temperatures in the temperate growing environment, conditions that are not encountered by any of the other subpopulations. On the other hand, the strong photosensitivity and late flowering observed in the *aromatic* varieties grown in the Himalayan foothills in northern India, Pakistan and Iran have been selected to enhance the highly valued grain quality and fragrance of accessions in this subpopulation.

Future Perspective in Flowering Time Studies

Flowering time is one of the most actively studied areas in plant science. It has

been the focus of basic research using model systems where it is featured as a fundamental biological and developmental process. It has also been intensively studied in the context of translational research in crop plants due to its impact on yield and adaptation. A review of the literature on flowering time in plants describes it as: 1) a process marking the transition from vegetative to reproductive growth that is highly conserved throughout the plant kingdom regardless of growth habit (determinant vs indeterminant) or life cycle (annual vs perennial) and that involves the coordination of endogenous developmental programs with exogenous environmental cues such as temperature, photoperiod, light quality and nutrient availability; 2) a process that is repressible by the manipulation of environmental cues such as photoperiod and temperature (particularly in the form of vernalization); 3) an evolutionary process whereby wild types within various species are late flowering and derived types are early, suggesting that late flowering is the ancient or ancestral state, while selection for early flowering is associated with artificial or human selection and systematic plant breeding to allow the expansion of crop cultivation into growing environments beyond the centers of origin and 4) a process in which variation is largely controlled by differences in gene expression and post-translational modification of a few gene products, in particular the mobile signal (florigen) that is produced by the *FLOWERING Time LOCUS T* (*FT*) gene in *Arabidopsis* and its orthologous counterparts in other species.

For the past two decades, forward genetics, reverse genetics and comparative genomics have dominated the research agenda in flowering time with the objective of understanding the basic process by which plants perceive and use environmental information to coordinate their biological growth processes. The science has progressed based on the cloning and functional analysis of genes that participate in the four known interacting pathways that determine flowering time in plants (the

photoperiod, vernalization, giberellic acid and autonomous pathways) and is expanding to include the identification of cis and trans-acting elements that interact with known flowering time genes and integrate environmental signals with the different flowering pathways. Although much is known about the function and potential interaction of genes in the different flowering time pathways, little is known about the protein-protein interactions involved in the transport of the HD3A/FT protein from the leaves through the phloem to the apical meristem, where it ultimately interacts with *FLOWERING LOCUS D* to change the cell fate from the vegetative to the reproductive phase. A better understanding of the mechanism that regulates *HD3A* expression, transport and interaction with other gene products is key to our understanding of flowering time and will greatly enhance our ability to manipulate it.

The exploration of experimental populations has had limited impact on the ability of plant breeders to manipulate the extensive pools of genetic variation present in cultivated species. This is due to a lack of understanding about how diverse alleles at multiple loci in the flowering time pathways interact among themselves and with the environment to condition the *days-to-flowering* phenotype. Breeders must deal with the further complexities of epistatic and pleiotropic interactions governing the relationship between flowering time, yield and grain quality in crop plants.

In our work, the use of the advanced backcross population between the *tropical japonica* US elite cultivar Jefferson and its wild relative *O. rufipogon* provided a useful starting point for capturing novel alleles related to flowering time in rice, and offered an opportunity to dissect the genetics underlying transgressive early flowering. Further exploration of flowering time using the rice diversity panel (500 accessions) that has recently been genotyped with 44,000 SNPs will allow the identification of novel alleles conferring early and late flowering in rice. In addition, the availability of chromosome segment substitution lines (CSSLs) developed from crosses between a

wild donor and a cultivated recurrent parent provide a systematic way to explore the breeding value of diverse wild ancestors and promise to expedite the identification of useful genetic variation for flowering time and yield in diverse environments.

APPENDIX

DEVELOPMENT OF SUB-INTROGRESSION LINES (SILS) CONTAINING *O. RUFIPOGON* INTROGRESSIONS AT *HD1, HD3A, RFT1, FT-L8* AND *FTL*

We aim to develop a set of near isogenic lines (NILs) containing introgressions from *O. rufipogon* (#105491) at *HD1, HD3A* and *RFT1* in the Jefferson backgound. These NILs will be evaluated for flowering time under short (10 h) and long (14 h) days and compared with the flowering time of the sub-introgression lines (SILs) described previously. Transgressive early flowering was observed in SILs carrying *O. rufipogon* alleles across the *dth1.1a* region on chromosome 1. The wild alleles at *dth1.1a* were associated with a trans-acting effect, enhancing the expression of *HD3A* and *RFT1* (located on chromosome 6) in the Jefferson SILs. In lines carrying an *O. rufipogon* introgression at both *dth1.1a* (*GI/SOC1/FT-L8)* and *HD3A/RFT1,* late-flowering was observed. This indicated that *O. rufipogon* alleles across the *dth1.1a* region were necessary for the transgressive early response but only if there were Jefferson alleles at *HD3A* and *RFT1*. When there were *O. rufipogon* alleles across the *HD3A-RFT1* region, they negated the effect of *O. rufipogon* alleles in the *GI/SOC1/FT-L8* introgression on chromosome 1 and led to late flowering.

We were unable to examine the effect on flowering time of *O. rufipogon* alleles in the chromosome 6 region in combination with Jefferson alleles at *dth1.1a* because we did not have NILs containing *O. rufipogon* introgressions across the *HD1-HD3A-RFT1* region alone. By constructing these NILs, we aim to test the hypothesis that *O. rufipogon* alleles in the *HD3A*/*RFT1* region inhibit the expression of early flowering. We hypothesize that the *O. rufipogon* alleles code for disfunctional proteins at *HD3A* and/or *RFT1* that cannot be upregulated by either *O. rufipogon* or Jefferson

alleles at *GI/SOC1/FT-L8* (in the *dth1.1a* region). Crosses between the newly developed NILs and existing SILs will offer a chance to examine G X G and epistasis in relation to flowering time when different combinations of *O. rufipogon* alleles are combined in the Jefferson background.

We also want to determine whether an *O. rufipogon* allele at *HD1* affects flowering time, either alone or in combination with *O. rufipogon* introgressions elsewhere in the genome. As one of the major photoperiod sensitivity loci in rice, we are interested in examining the effect of both *O. rufipogon* and Jefferson alleles at *HD1* and exploring how *HD1* interacts with other genes in the flowering time pathway to induce early or late flowering. Once we have NILs that each contain a single flowering time gene candidate from the *O. rufipogon* parent, we can make all possible combinations of crosses to allow us to determine which genes are involved and how they interact to determine the phenotype. The ability to create lines containing different combinations of Jefferson and *O. rufipogon* alleles will provide new insights into the control of flowering time in rice and into the genetics of transgressive variation for complex traits such as flowering time.

To further investigate the *dth1.1* QTL, se also aim to develop an NIL containing an *O. rufipogon* introgression at *FTL*, a flowering time gene that falls right in the middle of the *dth1.1* introgression and was not captured in any of the previously developed SILs. The availability of additional SILs containing smaller introgressions around each of the flowering time candidates will be helpful as we begin to examine expression differences in the SILs on a genome-wide basis. They will also allow us to explore how inbreds differ from hybrids in terms of gene expression and flowering time.

Materials and Methods

Plant Materials. One of the main objectives in the study is to develop a complete panel of NILs carrying *O. rufipogon* alleles in flowering time candidate genes on chromosome 6 (*HD1, HD3A, RFT1*, and *HD3A/RFT1*) and on chromosome 1 (*FT-L8* and *FTL*). Table 1 describes the NILs to be developed and the crosses to be made.

Appendix Table 5.1. Plant materials for gene expression analysis and selection of missing SILs.

Type	Line	FT*	Cross	PA Number
Parental Lines	Jefferson	Early	Self	RA8824
	O. rufipogon	Late	Self	IRGC105491
Pre-SIL	P9-84	Late	Self	PA_29851
	P15-62	Early	Self	PA_29852
	P13-67	Early	Self	PA_29872
SILs	SIL_GI/SOC1	Transgressive	Self	PA_29855
	SIL_SOC1/FT-L8	Transgressive	Self	PA_29859
	SIL_GI/SOC1/FT-L8	Transgressive	Self	PA_29856
New SILs to develop	SIL_FT-L8	NA	SIL_SOC1/FT-L8 x Jeff	PA_29858
	SIL_FTL	NA	P15-62 x Jeff	PA_29871
	SIL_HD1	NA	P13-67 x Jeff	PA_29868
	SIL_HD3A	NA	P9-84 x Jeff	PA_29863
	SIL_HD3A/RFT1	NA	P9-84 x Jeff	PA_29869
	SIL_RFT1	NA	P9-84 x Jeff	PA_29870

Growth Conditions. The plants should be grown for 20 days in greenhouse conditions to ensure vigorous growth. Plants may be transferred to growth chambers under short-day conditions (10 h) at 30 °C by day and 25 °C by night (14 h). Light is provided by fluorescent white light (400–700 nm, 100 μmol m−2 s−1) with 70% humidity. Flowering time is defined as the time when the first panicle has 10% anthers visible. Sixteen to 20 independent plants should be used to score flowering time.

Gene Expression Analysis. Leaves will be harvested from plants grown under short-day conditions at 35, 50 and 65 days after germination at the times corresponding to

the expression peak for each gene, as follows: 0-4 h (*HD3A* and *RFT1*), 8-12 h (*OsMADS50*), 12-16 h (*HD1*), 16-20 h (*OsGI*) and 20-24 h (*EHD1*). Leaf tissue is ground in liquid nitrogen using the genogrinder. Total RNA is extracted using Quiagen RNAesy Kit and treated with DNase I (Invitrogen). cDNA (20 µl) may be synthesized from 1 µg of total RNA using SuperScriptII Reverse Transcriptase (Invitrogen). One µl of cDNA is used for the real-time PCR performed with the SYBR Green PCR master mix (Applied Biosystems). Data are collected using the ABI PRISM 7000 sequence detection system. All expression levels are normalized by that of ubiquitin. Four to 6 independent plants should be used for this assay. Displayed data should represent means of 2 separate RNA extractions. Primers for gene expression analysis are summarized in Table 2.

Appendix Table 2. Primer sets for gene expression analysis.

Primer	Forward
OsGI-F	ATCGTTCTGCAGGCCGAGA
OsGI-R	TCACCAATGCTTCTGGGCTAT
HD1-F	TCAGCAACAGCATATCTTTCTCATCA
HD1-R	TCTGGAATTTGGCATATCTATCACC
HD3A-F	GCTCACTATCATCATCCAGCATG
HD3A-R	CCTTGCTCAGCTATTTAATTGCATAA
RFT1-F	TGGGTTAGCTGACCTAGATTCAAA
RFT1-R	GCCAACCACAAGAGGATCGT
OsSOC1-F	CAGGCCAGGAATAAGCTGGAT
OsSOC1-R	TTAGGATGGTTTGGTGTCATTGC
EHD1-F	GCGCTTTTGATTTCCTGC
EHD1-R	ATATGTGCTGCCAAATGTTGCT
UBQ-F	AACCAGCTGAGGCCCAAGA
UBQ-R	ACGATTGATTTAACCAGTCCATGA

Selection of new NILs. The study will require the selection of six new SILs: SIL_HD3A/RFT1, SIL_HD3A, SIL_RFT1, SIL_HD1, SIL_FT-L8 and SIL_FTL, to complete the series. The breeding strategy is presented in Figure 1 and includes three

pre-SIL lines (P9-84, P15-62 and P13-64) that can be backcrossed to Jefferson to generate the lines outlined above.

Optimization of *OsSOC1* Expression. The main constraint in the expression analysis of OsSOC1 is the specificity of the reaction. Although we were able to detect low RNA levels, a secondary peak was always present in the dissociation curves of the RT-PCR reaction indicating the presence of a secondary product that could affect the interpretation of results. To overcome this we tested several variables including:

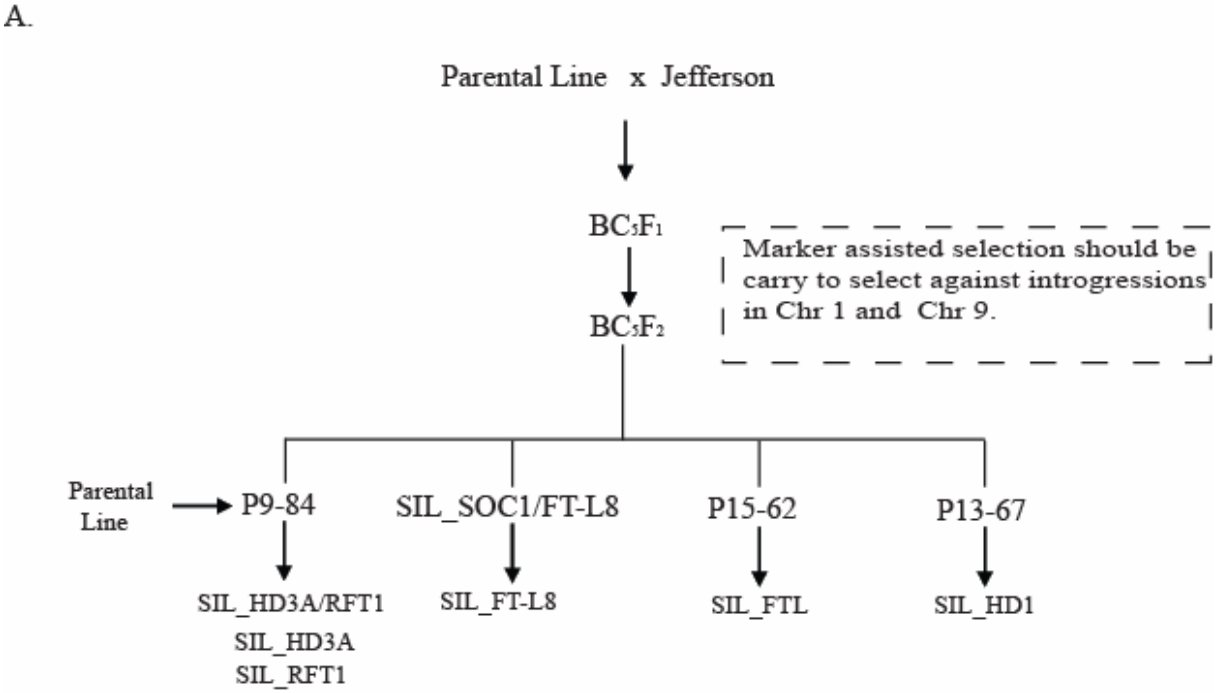
- Annealing Temperature: 50, 55 and 57 ^{o}C.
- cDNA concentration: 1, 3 and 5 ug/ul.
- Number cycles: 20, 30 and 35 amplification cycles. Increasing the cycles beyond 35 increases the chances of non-specific product amplification.
- Primer concentration: 1 and 2 mM/reaction.

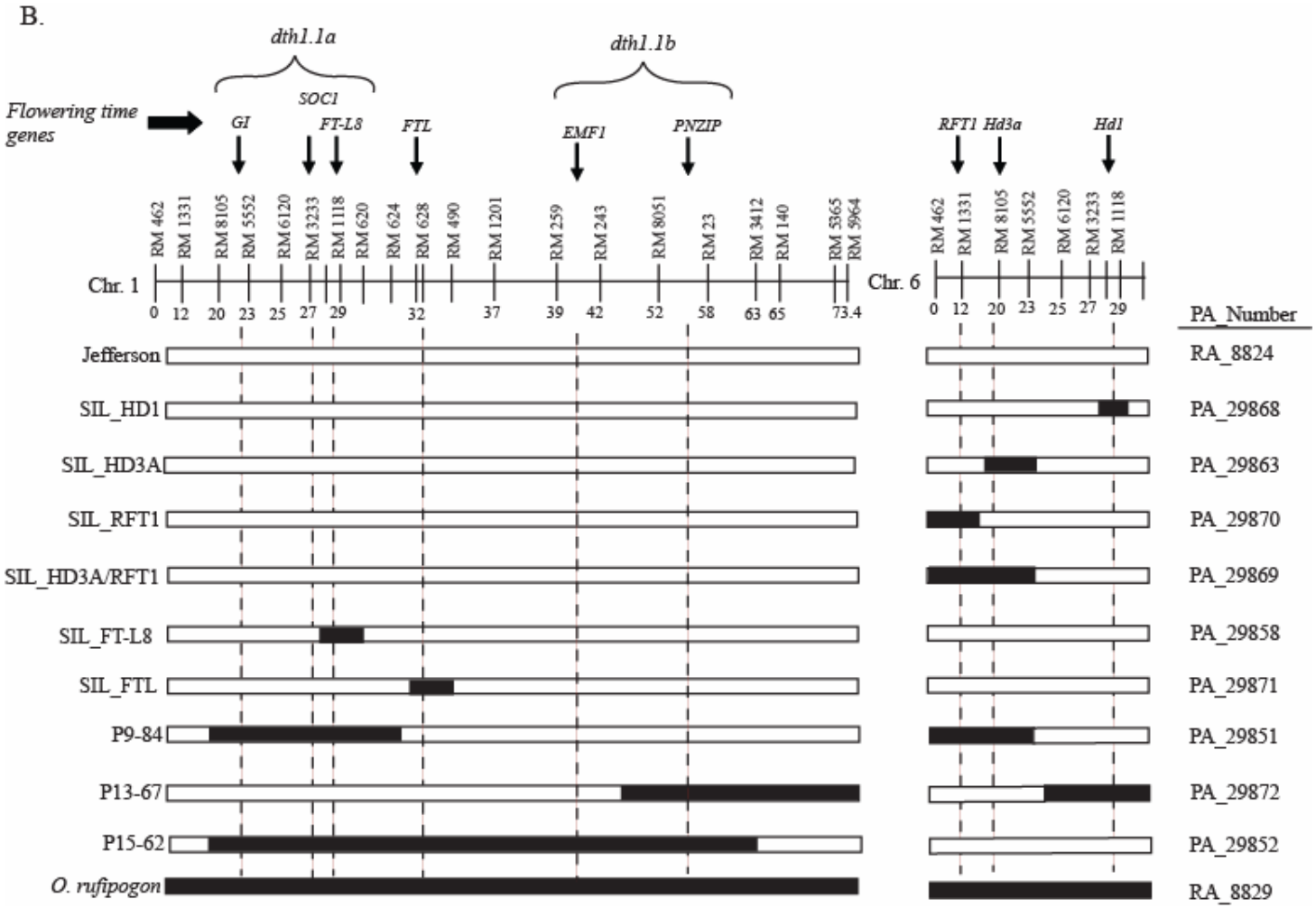
In the future, it may be necessary to redesign the primers to obtain useful RT-PCR results.

The importance of being able to track expression levels of *SOC1* in the NILs is to test the hypothesis that *SOC1* expression is correlated with both flowering time and expression of *HD3A.* The most interesting hypothesis is that the *O. rufipogon* allele at *SOC1* may drive the expression of *HD3A,* and together, the expression level(s) of these proteins would determine early flowering. Alternatively there may be other genes, closely linked to *SOC1,* that are responsible for the transgressive early flowering response in the SILs.

Appendix Figure 5.1. (A) Breeding scheme for development of SIL_FTL, SIL_FT-L8, SIL_HD3A/RFT1, SIL_RFT1, SIL_HD3A and SIL_HD1. (B) Graphical genotypes of the SILs and parental lines showing regions of *O. rufipogon* introgression (black rectangle= homozygous) across the *dth1.1a* region of chromosome 1 and the *RFT1/HD3A/HD1* region of chromosome 6. Position of candidate flowering time genes indicated by vertical arrows across top in relation to SSR markers.

A.


Parental Line x Jefferson
BC3F1
BC3F2
Marker assisted selection should be carry to select against introgressions in Chr 1 and Chr 9.
Parental Line
P9-84
SIL_SOC1/FT-L8
P15-62
P13-67
SIL_HD3A/RFT1
SIL_HD3A
SIL_RFT1
SIL_FT-L8
SIL_FTL
SIL_HD1


B.
dth1.1a
dth1.1b
Flowering time genes
SOC1
GI
FT-L8
FTL
EMF1
PNZIP
RFT1
Hd3a
Hd1
RM 462
RM 1331
RM 8105
RM 5552
RM 6120
RM 3233
RM 1118
RM 620
RM 624
RM 628
RM 490
RM 1201
RM 259
RM 243
RM 8051
RM 23
RM 3412
RM 140
RM 5365
RM 5964
Chr. 1
0 12 20 23 25 27 29 32 37 39 42 52 58 63 65 73.4
Chr. 6
0 12 20 23 25 27 29
PA_Number
Jefferson
SIL_HD1
SIL_HD3A
SIL_RFT1
SIL_HD3A/RFT1
SIL_FT-L8
SIL_FTL
P9-84
P13-67
P15-62
O. rufipogon
RA_8824
PA_29868
PA_29863
PA_29870
PA_29869
PA_29858
PA_29871
PA_29851
PA_29872
PA_29852
RA_8829

REFERENCES

Kim SL, Lee S, Kim HJ, Nam HG, An G (2007) OsMADS51 is a short-day flowering promoter that functions upstream of EHD1, OsMADS14, and HD3A. Plant Physiol 145:1484-1494.

Komiya R, Yokoi S, Shimamoto K (2009) A gene network for long-day flowering activates RFT1 encoding a mobile flowering signal in rice. Development 136: 3443-3450.

Komiya R, Yokoi S, Shimamoto K (2008) Hd3a and RFT1 are essential for flowering in rice. Development 135: 767-774.

Lee, S., Kim, J., Han, J. J., Han, M. J. and An, G. (2004). Functional analyses of the flowering time gene OsMADS50, the putative SUPPRESSOR OF OVEREXPRESSION
OF CO 1/AGAMOUS-LIKE 20 (SOC1/AGL20) ortholog in rice. Plant J. 38, 754-764.

Ryu CH, Lee S, Cho LH, et al. (2009) OsMADS50 and OsMADS56 function antagonistically in regulating long day (LD)-dependent flowering in rice. Plant cell and environment 32: 1412-1427.

Takahashi Y, Teshima KM, Yokoi S, Innan H, Shimamoto K (2009) Variations in HD1 proteins, HD3A promoters, and EHD1 expression levels contribute to diversity of flowering time in cultivated rice. Proc Natl Acad Sci USA 106:4555-4560.